

FEATURES	Location/Qualifiers
source	1..8589
	/organism="unknown"
ORIGIN	/mol_type="unassigned DNA"
Query Match	97.5%; Score 4184.8; DB 6; Length 8589;
Best Local Similarity	99.7%; Pred. No. 0;
Matches 4232; Conservative	0; Mismatches 47; Indels 9; Gaps 1;
QY	3 GGTGGCCGACGCGCGCAACTCAGGCAAGGCTGAGGAGGAGCGCGCGGCGGCGAGC 62
DB	9 GGTGGCCGACGCGCGCAACTCAGGCAAGGCTGAGGAGGAGCGCGCGGCGGCGAGC 68
QY	63 GCTCGGGGACGAGCGCGGCGCGCACTGCGGACTGGGACGTGACACGAGGCTGGGAGGCC 122
DB	69 GCTCGGGGACGAGCGCGGCGCGCACCGAACCCTGCAACCGCGACAGGGGCGCTGGAGGCC 128
QY	123 GGGGCTCGGGGCGCGGCTCGGCTCCCGGGCTGCTCTCACCGTGGGGCCACGGCT 182
DB	129 GGGACCGCGCGCGGCTGCTCTCCCGGGGCTGCTCGCGGGGCGCTGCCCCCGCGCC 188
QY	183 GCTGCTGCTGCTGTTGCTCCCGCGCGGCTGTTGCTGCTGCTGCTGCTGCTGAGGC 242
DB	189 GCTGCTGCTGCTGCTCTT-----TTGCTGCTGCTGCTGCTGCTGCTGCGGAGGC 239
QY	243 CGAGGCGCGGCGCGCGCGGCTGTCGGGCTCAGCCGACGCGGCGGCAAGGAATG 302
DB	240 CGAGGCGGCTGCGGCTGGCGGCGGCGGCTGTCGGGCTCGCGCGGCGGCGGCGGAGG 299
QY	303 TGACCGGCGCTGCTGTCACGCGGCGTGTGCAACCCCTGGCACCGCGGCGAGTGCCTGCCCC 362
DB	300 TGACCGGCGCTGCTGTCACGCGGCGTGTGCAACCCCTGGCACCGCGGCGAGTGCCTGCCCC 359
QY	363 CGCGGCTGGGTGGCGGACGACATGCGACGCTGCGGGGCGCGCTTCAGACTGATC 422
DB	360 CGCGGCTGGGTGGCGGACGACATGCGACGCTGCGGGGCGCGCTTCAGACTGATC 419
QY	423 TTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAATACAAACGAAGTGACGTTGGCT 482
DB	420 TTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAATACAAACGAAGTGACGTTGGCT 479
QY	483 CATTGAAGACAGCCAAATAGAAATAGACACTTCGTTTCAATCAATTTGCTACAGAGTG 542
DB	480 CATTGAAGACAGCCAAATAGAAATAGACACTTCGTTTCAATCAATTTGCTACAGAGTG 539
QY	543 TAGTTGGGACCAATTTATATGTTATGATGGGACTCAATTTATGCAACGCTAGTCTGCTGC 602
DB	540 TAGTTGGGACCAATTTATATGTTATGATGGGACTCAATTTATGCAACGCTAGTCTGCTGC 599
QY	603 ATTAGTGGCTCATGTTTCCTGAGAGAGATGGCAATGACACTGCTCCCTGAGGTTTGTGC 662
DB	600 ATTAGTGGCTCATGTTTCCTGAGAGAGATGGCAATGACACTGCTCCCTGAGGTTTGTGC 659
QY	663 CACATCAGGTATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGACTGGATT 722
DB	660 CACATCAGGTATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGACTGGATT 719
QY	723 TAATATTACTTACAGTTTGTATGTTGCTCAAAATAACTGCTCAGGCGCGGAGAGGTAA 782
DB	720 TAATATTACTTACAGTTTGTATGTTGCTCAAAATAACTGCTCAGGCGCGGAGAGGTAA 779
QY	783 GATCAGTAATAGCAGGAACTGTTGATGTGAATGTTCTGAAATGTTCTGAAACTGGAAAGGTGAAGC 842
DB	780 GATCAGTAATAGCAGGAACTGTTGATGTGAATGTTCTGAAATGTTCTGAAACTGGAAAGGTGAAGC 839
QY	843 ATGTGACATTCCTCAGTCTGACAGCAACTGTGGTTTTTCTTCATCGAGGCACTCTGCAATTC 902
DB	840 ATGTGACATTCCTCAGTCTGACAGCAACTGTGGTTTTTCTTCATCGAGGCACTCTGCAATTC 899
QY	903 AAGTGATGTCAGAGATGCTCCTGCTTCTCAGACTGGCAGGTCCTGGAATGTTTCAGTTCC 962
DB	900 AAGTGATGTCAGAGATGCTCCTGCTTCTCAGACTGGCAGGTCCTGGAATGTTTCAGTTCC 959

Qy	2043	AACGTATGAACAAGAAAAAGTTTAAAAATCAGAAATGTTTTTCCAAAGAAACTCTTGACCA	2107
Db	2040	AACGTATGAACAAGAAAAAGTTTAAAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCA	2099
Qy	2103	TGACAGATGTGACCCAGCACACAGATTGTTACAGCTGTACAGTGCACCAACCAATCACTGGCCA	2162
Db	2100	TGACAGATGTGACCCAGCACACAGATTGTTACAGCTGCACAGCCCAACCAATGACTGGCA	2159
Qy	2163	CTGGTGAANTGACCAATTGTGTCCACAGGAACCAAGCTGCTCAGAGGCCAGATCTCCAT	2222
Db	2160	CTGGTGAANTGACCAATTGTGTCCCCAGGAACCAAGCTGCTCAGAGGCCAGATCTCCAT	2219
Qy	2223	TTTTAGGTATGAGAAATTTGCCCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG	2282
Db	2220	TTTTAGGTATGAGAAATTTGCCCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG	2279
Qy	2283	CTGCAGAGCTGTGCCCTGTGCACAGAACTGCAGTGGAGCCCGGAATCAGAGATGCAT	2342
Db	2280	CTGCAGAGCTGTGCCCTGTGCACAGAACTGCAGTGGAGCCCGGAATCAGAGATGCAT	2339
Qy	2343	TGCCCTGCCGAAAATACTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAA	2402
Db	2340	TGCCCTGCCGAAAATACTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAA	2399
Qy	2403	AATTACTACTGCCAAGGAGAAATTATGACAAATGCTAAATGCTCTGTAGGAACACAAATGC	2462
Db	2400	AATTACTACTGCCAAGGAGAAATTATGACAAATGCTAAATGCTCTGTAGGAACACAAATGC	2459
Qy	2463	CCTTTGGCTTCTCTTCAACCCAGAAAGAGTAGAAATTTGTCTTAAAGCAGCTGCGAAT	2522
Db	2460	CCTTTGGCTTCTCTTCAACCCAGAAAGAGTAGAAATTTGTCTTAAAGCAGCTGCGAAT	2519
Qy	2523	AATGCAGTCACTCAGAGCATGTCCAAGCTCACTTAAACCCCATGGTGGCCCTTCGGAA	2582
Db	2520	AATGCAGTCACTCAGAGCATGTCCAAGCTCACTTAAACCCCATGGTGGCCCTTCGGAA	2579
Qy	2583	GATCAATGTGTCTTACTGTGGTGGGAAGATATGTCCCCATTTACAAATAGTTTATACAA	2642
Db	2580	GATCAATGTGTCTTACTGTGGTGGGAAGATATGTCCCCATTTACAAATAGTTTATACAA	2639
Qy	2643	GTGGATGCCCTCTGAGCCCAAGTGAATGCTGGAATTTTTCAGAAACCCAGTAC	2702
Db	2640	GTGGATGCCCTCTGAGCCCAAGTGAATGCTGGAATTTTTCAGAAACCCAGTAC	2699
Qy	2703	TGCGGGACTGAAGGCTGCAACCTGCATCAACCACTCAATGTTAGTGTCTGTGAAAGGCC	2762
Db	2700	TGCGGGACTGAAGGCTGCAACCTGCATCAACCACTCAATGTTAGTGTCTGTGAAAGGCC	2759
Qy	2763	TGCAAAACACAGTCTAAGCAGTGCAGGACACCAATGTGCTTCAAGGACAGCATGTGGAGA	2822
Db	2760	TGCAAAACACAGTCTAAGCAGTGCAGGACACCAATGTGCTTCAAGGACAGCATGTGGAGA	2819
Qy	2823	TTGCACAGGGGCAAGCTCTGAGTGCATGTGGTGGAGCAACAAGACAGTGTGTGACTC	2882
Db	2820	TTGCACAGGGGCAAGCTCTGAGTGCATGTGGTGGAGCAACAAGACAGTGTGTGACTC	2879
Qy	2883	CAATGCCCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC	2942
Db	2880	CAATGCCCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC	2939
Qy	2943	CTGCCCCCTGAAAATTTGTTTCAGGCTACTGTACCTGTAGTCATTGCTTGAGGCAACGAGG	3002
Db	2940	CTGCCCCCTGAAAATTTGTTTCAGGCTACTGTACCTGTAGTCATTGCTTGAGGCAACGAGG	2999
Qy	3003	CTGTGGGTGTGTACTGATCCAGCAATACTGGCAAGGGAAATGCTATAGAGGTTCTCTTA	3062
Db	3000	CTGTGGGTGTGTACTGATCCAGCAATACTGGCAAGGGAAATGCTATAGAGGTTCTCTTA	3059
Qy	3063	TAAAGGACCAAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTATCCACAGCCCT	3122
Db	3060	TAAAGGACCAAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTATCCACAGCCCT	3119
Qy	3123	GCTCAATTCACGACA TGTGTCTAGAGGACAGCAGATACAACCTGGTCTTTTCAATCACTGTCC	3182

1081 TTCAACCACTCAGATATTAACAATGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTT 1140 Db
1141 CCACTAAACCGTCTCTGTGAACAATGTTGGTTGTAGATATGGTCATCTTTTGGCAATATAC 1200 Qy
1141 CCACTAAACCGTCTCTGTGAACAATGTTGGTTGTAGATATGGTCATCTTTTGGCAATATAC 1200 Db
1201 AAGGATAAATTTACATGATGAGGAGAAATTTGATCCCACTGGGAATGTGACCAATGAG 1260 Qy
1201 AAGGATAAATTTACATGATGAGGAGAAATTTGATCCCACTGGGAATGTGACCAATGAG 1260 Db
1261 TTGAGAGTCTTTTACATATGATGATGAGTCTTGTGACCCCTTAAGGCAAAAGGAG 1320 Qy
1261 TTGAGAGTCTTTTACATATGATGATGAGTCTTGTGACCCCTTAAGGCAAAAGGAG 1320 Db
1321 CAGTATGCAATGTTGGGCACTCTGCACACATTTGTACCTGAAAGATGGCCGAGTGTCT 1380 Qy
1321 CAGTATGCAATGTTGGGCACTCTGCACACATTTGTACCTGAAAGATGGCCGAGTGTCT 1380 Db
1381 ATGCTGTCTCATCTTTTGTGCTCACTGCTCTCTATGGATATATAAGCAATGTGCAAGATAT 1440 Qy
1381 ATGCTGTCTCATCTTTTGTGCTCACTGCTCTCTATGGATATATAAGCAATGTGCAAGATAT 1440 Db
1441 GATTTGGATTAAGAACACATGAGTATATTACACACCCAGGTTGCCCTTGTGCAAGGGGT 1500 Qy
1441 GATTTGGATTAAGAACACATGAGTATATTACACACCCAGGTTGCCCTTGTGCAAGGGGT 1500 Db
1501 TAGCGGCATAGCAGTGTTTACGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTAC 1560 Qy
1501 TAGCGGCATAGCAGTGTTTACGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTAC 1560 Db
1561 AAGGCTTTAGTGGCAATTAAGTACCGGCTTGCAATGATCTCTACGGATATGATGGAT 1620 Qy
1561 AAGGCTTTAGTGGCAATTAAGTACCGGCTTGCAATGATCTCTACGGATATGATGGAT 1620 Db
1621 ACCCAGATGAGGACCACTTTTAAGGACAGCCGATTTTCCGTTACTTGGCACACAGCTGTG 1680 Qy
1621 ACCCAGATGAGGACCACTTTTAAGGACAGCCGATTTTCCGTTACTTGGCACACAGCTGTG 1680 Db
1681 ATAGTGAAGTGAACCAATGCTGTGTTGGGGGAAACACACAAATGAACATCTATGAGC 1740 Qy
1681 ATAGTGAAGTGAACCAATGCTGTGTTGGGGGAAACACACAAATGAACATCTATGAGC 1740 Db
1741 CATGGGCGCAAAATGCTTCTTCCAGATTTCAATGGCCTATGACATTTGCCCTGTGACCGCTGG 1800 Qy
1741 CATGGGCGCAAAATGCTTCTTCCAGATTTCAATGGCCTATGACATTTGCCCTGTGACCGCTGG 1800 Db
1801 TCAGTGTCTCCAGACCTGATCTCCAACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860 Qy
1801 TCAGTGTCTCCAGACCTGATCTCCAACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860 Db
1861 TTACACAAAGCAGCACCATGATGTTGGTGGTTTCAATGATCTCTCTCCTCAGCGACATC 1920 Qy
1861 TTACACAAAGCAGCACCATGATGTTGGTGGTTTCAATGATCTCTCTCCTCAGCGACATC 1920 Db
1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATTCGGAGTGAAGCGCTTGTATACAGCA 1980 Qy
1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATTCGGAGTGAAGCGCTTGTATACAGCA 1980 Db
1981 GGACCTGGTATTCGGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCTGGGGCGGTG 2040 Qy
1981 GGACCTGGTATTCGGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCTGGGGCGGTG 2040 Db
2041 GCAACTGATGAACAAGAGAAAGTAAATCAGATGTTTTCGAAAGAACTCTTGAC 2100 Qy
2041 GCAACTGATGAACAAGAGAAAGTAAATCAGATGTTTTCGAAAGAACTCTTGAC 2100 Db
2101 CATGACAGATGTGACAGCAGACAGATGTTTACAGCTGTACAGCCAAACCAATCAGCTGC 2160 Qy
2101 CATGACAGATGTGACAGCAGACAGATGTTTACAGCTGTACAGCCAAACCAATCAGCTGC 2160 Db
2161 CACTGTGCAATGACCAATGTTGTCCCGAGAACCAAGCTGTCTGAGAGGCGCAGATCTCC 2220 Qy
2161 CACTGTGCAATGACCAATGTTGTCCCGAGAACCAAGCTGTCTGAGAGGCGCAGATCTCC 2220 Db

2221 ATTTTATAGTATGAGAAATGCCCCAAGGATACCCCCATGTACTACTGTAAACAAGAGACC 2280 Qy
2221 ATTTTATAGTATGAGAAATGCCCCAAGGATACCCCCAAGGATACCCCCAAGAGACC 2280 Db
2281 AGCTGCAGAGCTGTGCCCTGACAGCACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2340 Qy
2281 AGCTGCAGAGCTGTGCCCTGACAGCACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2340 Db
2341 ATTGCCCTCCCGAAATATCTGTGCAATGCTGCTGGCAATTTGGTTGGAAATCTCATGTTTG 2400 Qy
2341 ATTGCCCTCCCGAAATATCTGTGCAATGCTGCTGGCAATTTGGTTGGAAATCTCATGTTTG 2400 Db
2401 AATAATTTACTGCTCCAGAGCAATTTATGACATGCTAAATTTGTTCTGTAGGAACCAAT 2460 Qy
2401 AATAATTTACTGCTCCAGAGCAATTTATGACATGCTAAATTTGTTCTGTAGGAACCAAT 2460 Db
2461 GCCCTTTTGGCTTCTTTACAAACCCAGAGAGTAGAATTTGCTTTAAGCAGCTGCGA 2520 Qy
2461 GCCCTTTTGGCTTCTTTACAAACCCAGAGAGTAGAATTTGCTTTAAGCAGCTGCGA 2520 Db
2521 ATAAATGCACTCATCTCAGAGCATGCTCCAGCTCACCTTAAACCCATGGTGGCTTCGG 2580 Qy
2521 ATAAATGCACTCATCTCAGAGCATGCTCCAGCTCACCTTAAACCCATGGTGGCTTCGG 2580 Db
2581 AAGATCAATGTGTCTTACTGCTGGGAGATATGTCCTCCATTTACAAATAGTTTACTA 2640 Qy
2581 AAGATCAATGTGTCTTACTGCTGGGAGATATGTCCTCCATTTACAAATAGTTTACTA 2640 Db
2641 CAGTGAATGCCCTGAGCCAGTATGCTGATGCTGGAATTTTATCAGAACCCAGT 2700 Qy
2641 CAGTGAATGCCCTGAGCCAGTATGCTGATGCTGGAATTTTATCAGAACCCAGT 2700 Db
2701 ACTCGGGGACTGAAGGCTCAACCTGCAATCAACCCCAATGCTGTGTAAGAG 2760 Qy
2701 ACTCGGGGACTGAAGGCTCAACCTGCAATCAACCCCAATGCTGTGTAAGAG 2760 Db
2761 CTTGCAAAACACAGTCTTAAGCAGTCCCGACACCATGCTGTCAGACAGCATGTGGA 2820 Qy
2761 CTTGCAAAACACAGTCTTAAGCAGTCCCGACACCATGCTGTCAGACAGCATGTGGA 2820 Db
2821 GATTGACACAGGGGAGCTCTGAGTGCATGCTGGTGCAGCAACATGAAGCAGTGTGTGAC 2880 Qy
2821 GATTGACACAGGGGAGCTCTGAGTGCATGCTGGTGCAGCAACATGAAGCAGTGTGTGAC 2880 Db
2881 TCCAAATGCCATGTGGCTCCCTTCCCTTTGGCCAGTGTATGGAATGATACAGTACG 2940 Qy
2881 TCCAAATGCCATGTGGCTCCCTTCCCTTTGGCCAGTGTATGGAATGATACAGTACG 2940 Db
2941 ACCTGCCCCCTGAAATTTGTTTCCAGCTACTGCTAGTCTAGTCTTGGAGCAACA 3000 Qy
2941 ACCTGCCCCCTGAAATTTGTTTCCAGCTACTGCTAGTCTAGTCTTGGAGCAACA 3000 Db
3001 GGCTGTGGCTGTGATGATCCAGCAATATGCAAGAGGAATGATAGAGGTTCC 3060 Qy
3001 GGCTGTGGCTGTGATGATCCAGCAATATGCAAGAGGAATGATAGAGGTTCC 3060 Db
3061 TATAAAGGACCAAGTGAAGATGCTTCCGAAAGCCCTACAGGAAATTTCTATCCACAGCC 3120 Qy
3061 TATAAAGGACCAAGTGAAGATGCTTCCGAAAGCCCTACAGGAAATTTCTATCCACAGCC 3120 Db
3121 CTGCTCAATTCAGATGTTCTAGAGGACAGCATCACTGTCTTTTCACTGAT 3180 Qy
3121 CTGCTCAATTCAGATGTTCTAGAGGACAGCATCACTGTCTTTTCACTGAT 3180 Db
3181 CCAGCTTGCAATGCAACCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240 Qy
3181 CCAGCTTGCAATGCAACCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240 Db
3241 GAGAACTGACACAGGGAAGCATTGCGAGACCTGATATCTGGCTTCTACGGTATCCC 3300 Qy
3241 GAGAACTGACACAGGGAAGCATTGCGAGACCTGATATCTGGCTTCTACGGTATCCC 3300 Db

QY 1263 GAGAGTTTTCACATTCATATGAGTCATGGTGTGTTGACCCCTAAGCCAAAGAGCA 1322
Db 1260 GAGAGTTTTCACATTCATATGAGTCATGGTGTGTTGACCCCTAAGCCAAAGAGCA 1319
QY 1323 GTATGCAAGTGTGGGCACTCTGCACACATTGTTACACTGAAGAAATGGCGAGTGGTCAT 1382
Db 1320 GTATGCAAGTGTGGGCACTCTGCACACATTGTTACACTGAAGAAATGGCGAGTGGTCAT 1379
QY 1383 GCTGGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGAAATAGA 1442
Db 1380 GCTGGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGAAATAGA 1439
QY 1443 TTTGCAATAAGAACACATGAGAGTATATACACACCCAGGGTGCCCTGTGCAGAGGGGTTA 1502
Db 1440 TTTGCAATAAGAACACATGAGAGTATATACACACCCAGGGTGCCCTGTGCAGAGGGGTTA 1499
QY 1503 CGGCCATAGCAGTGTATACACCATAGGACCAAGGCCCTTATACGTTTCATGGTGGCTACAA 1562
Db 1500 CGGCCATAGCAGTGTATACACCATAGGACCAAGGCCCTTATACGTTTCATGGTGGCTACAA 1559
QY 1563 GCTTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGATAC 1622
Db 1560 GCTTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGATAC 1619
QY 1623 CCAGATGGACCAATCTTAAGGACACCGCAATTTTCCGTTACTTGCACACAGCTGTGAT 1682
Db 1620 CCAGATGGACCAATCTTAAGGACACCGCAATTTTCCGTTACTTGCACACAGCTGTGAT 1679
QY 1683 AGTGAAGTGAACCAATGCTGTGTGTTGGGGAACACACACATGACACATCTATGAGCA 1742
Db 1680 AGTGAAGTGAACCAATGCTGTGTGTTGGGGAACACACACATGACACATCTATGAGCA 1739
QY 1743 TGGCGCCAAATGCTTCTTTCAGATTTTCATGGCCTATGACATTCGCTGTGACCGTGGTC 1802
Db 1740 TGGCGCCAAATGCTTCTTTCAGATTTTCATGGCCTATGACATTCGCTGTGACCGTGGTC 1799
QY 1803 AGTGTCTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGCTCT 1862
Db 1800 AGTGTCTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGCTCT 1859
QY 1863 ACACAGACGACCATGATGTGTTGGTGTGTTCAATAGTCTCTCTCCTCAGCGACATCT 1922
Db 1860 ACACAGACGACCATGATGTGTTGGTGTGTTCAATAGTCTCTCTCCTCAGCGACATCT 1919
QY 1923 GGTATTCACCTCGGAACAGTGTGATGGCATCGGAGTGAAGCGCTGTTAGCAGCAGG 1982
Db 1920 GGTATTCACCTCGGAACAGTGTGATGGCATCGGAGTGAAGCGCTGTTAGCAGCAGG 1979
QY 1983 ACCTGGTATTCGGTGTGTGGAAACAAGGCTCGTCTCAGTGTATCTCTGGGCGCTGC 2042
Db 1980 ACCTGGTATTCGGTGTGTGGAAACAAGGCTCGTCTCAGTGTATCTCTGGGCGCTGC 2039
QY 2043 AACTGATGAACAAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2102
Db 2040 AACTGATGAACAAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2099
QY 2103 TGACAGATGTGACCAAGACACAGATGTTTACAGCTGTACAGCCAAACACCAATGACTGCCA 2162
Db 2100 TGACAGATGTGACCAAGACACAGATGTTTACAGCTGTACAGCCAAACACCAATGACTGCCA 2159
QY 2163 CTGGTGAATGAACATTTGTGTCCCGAGAACACAGCTGCTCAGAGGCCAGATCTCAT 2222
Db 2160 CTGGTGAATGAACATTTGTGTCCCGAGAACACAGCTGCTCAGAGGCCAGATCTCAT 2219
QY 2223 TTTTATGATGAGAAATGGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG 2282
Db 2220 TTTTATGATGAGAAATGGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG 2279
QY 2283 CTGAGAGCTGTGCCCTTGACACCAAGAACTGCCAGTGGAGCCCGGAAATCAGAGTGCAT 2342
Db 2280 CTGAGAGCTGTGCCCTTGACACCAAGAACTGCCAGTGGAGCCCGGAAATCAGAGTGCAT 2339
QY 2343 TGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATCTCATGTTTGA 2402

Db 2340 TGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATTTGGTGGCAATCTCATGTTTGA 2399
QY 2403 AATTACTACTGCCAAGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGAAACCAATGC 2462
Db 2400 AATTACTACTGCCAAGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGAAACCAATGC 2459
QY 2463 CTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCACTGCGAAT 2522
Db 2460 CTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCACTGCGAAT 2519
QY 2523 AATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTTAACCCCATGGCTGGGCTTCGAA 2582
Db 2520 AATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTTAACCCCATGGCTGGGCTTCGAA 2579
QY 2583 GATCAATGTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCACTGCGAAT 2642
Db 2580 GATCAATGTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCACTGCGAAT 2639
QY 2643 GTGATGSCCTCTCAGCCAGTCATGCTGGATTTCTGTGAAATTTTATCAGAACCCAGTAC 2702
Db 2640 GTGATGSCCTCTCAGCCAGTCATGCTGGATTTCTGTGAAATTTTATCAGAACCCAGTAC 2699
QY 2703 TCGGGGAATGAAGGCTGCAACCTGTCATCAACCCCATCAATGGTGTGTTCTGTGAAGGCC 2762
Db 2700 TCGGGGAATGAAGGCTGCAACCTGTCATCAACCCCATCAATGGTGTGTTCTGTGAAGGCC 2759
QY 2763 TGCAAAACCAAGTCTAAAGCAGTGCAGACACCATGTGCTTGGAGACAGCATGTGGAGA 2822
Db 2760 TGCAAAACCAAGTCTAAAGCAGTGCAGACACCATGTGCTTGGAGACAGCATGTGGAGA 2819
QY 2823 TTGCAACAGCGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAAGTGTGGACTC 2882
Db 2820 TTGCAACAGCGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAAGTGTGGACTC 2879
QY 2883 CAATGCTATGTGCGCTCTTCCCTTTGSCAGTGTATGGAATGCTATACCATGAGCAC 2942
Db 2880 CAATGCTATGTGCGCTCTTCCCTTTGSCAGTGTATGGAATGCTATACCATGAGCAC 2939
QY 2943 CTGCCCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCCAG 3002
Db 2940 CTGCCCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCCAG 2999
QY 3003 CTGTGGCTGTGTACTGATCCAGCAATACTGCAAGAGGAAATGATCATAGAGGTTCTTA 3062
Db 3000 CTGTGGCTGTGTACTGATCCAGCAATACTGCAAGAGGAAATGATCATAGAGGTTCTTA 3059
QY 3063 TAAAGGACCAAGTGAAGATGCTTTCAGACCCCTTACAGAAATTTCTATPCACAGCCCT 3122
Db 3060 TAAAGGACCAAGTGAAGATGCTTTCAGACCCCTTACAGAAATTTCTATPCACAGCCCT 3119
QY 3123 GCTCAATTCAGCATGTGTAGAGACAGACATACACTGCTGCTTTTCATTCATGCTCC 3182
Db 3120 GCTCAATTCAGCATGTGTAGAGACAGACATACACTGCTGCTTTTCATTCATGCTCC 3179
QY 3183 AGCTTGCATATGCAAGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGA 3242
Db 3180 AGCTTGCATATGCAAGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGA 3239
QY 3243 GAACCTGACACAGCAAGCACTGCGAGACCTGCAATGCTGCTTCTAGGTTGCTCCAC 3302
Db 3240 GAACCTGACACAGCAAGCACTGCGAGACCTGCAATGCTGCTTCTAGGTTGCTCCAC 3299
QY 3303 CAATGAGGGAATGTGACGATGCAAGTGAATGGGCAACGCTCTCTGTGCAACACCA 3362
Db 3300 CAATGAGGGAATGTGACGATGCAAGTGAATGGGCAACGCTCTCTGTGCAACACCA 3359
QY 3363 CAGGGCAAGTGTCTTGCACACCAAGCGCTCAGGGGAGCGAGTGCAGCTATGTA 3422
Db 3360 CAGGGCAAGTGTCTTGCACACCAAGCGCTCAGGGGAGCGAGTGCAGCTATGTA 3419
QY 3423 GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATTACTCTTCTTAT 3482

Db 3420 GGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATATCTTCTTAT 3479

Qy 3483 TGAATATCAGTTACCTTTAGTCTATCCAGAGAGATGATCGCTATTATACACAGCTACAA 3542

Db 3480 TGAATATCAGTTACCTTTAGTCTATCCAGAGAGATGATCGCTATTATACACAGCTACAA 3539

Qy 3543 TTTTGTGCTACTCTCTGACGAAACAAACAGAGATTGGAATGTTTCTATCAATCCCTCCAA 3602

Db 3540 TTTTGTGCTACTCTCTGACGAAACAAACAGAGATTGGAATGTTTCTATCAATCCCTCCAA 3599

Qy 3603 GAATTTCAACCTCAACATCACTGGCTGCGAGTTCTCAGCTGGACCCAGCTCGAGA 3662

Db 3600 GAATTTCAACCTCAACATCACTGGCTGCGAGTTCTCAGCTGGACCCAGCTCGAGA 3659

Qy 3663 AGAGATCCCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3722

Db 3660 AGAGATCCCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3719

Qy 3723 GAAGTTTGATTTGCGAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTAC 3782

Db 3720 GAAGTTTGATTTGCGAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTAC 3779

Qy 3783 CTGCCCCATCAAAATTCAGAT 3803

Db 3780 CTGCCCCATCAAAATTCAGT 3800

RESULT 4

AF531101 LOCUS 4493 bp mRNA linear MAM 19-AUG-2002

DEFINITION Bos taurus attractin mRNA, complete cds.

ACCESSION AF531101

VERSION AF531101.1 GI:22297307

KEYWORDS

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 4493)

Graphodatskaya, D., Joerg, H. and Stranzinger, G.

Direct submission

Submitted (18-JUN-2002) Animal Sciences, ETH, Tannenstrasse 1, Zurich, CH 8092, Switzerland

FEATURES

Location/Qualifiers

1..4493

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/chromosome="13"

52..4299

/note="transmembrane form; mahogany protein"

/codon_start=1

/product="attractin"

/protein_id="AA095445.1"

/db_xref="GI:22297308"

CDS

/translation="MVAATEARLRIRIVTTPAPTSRNGRRCDWDATRAGRPLRTGL
RLPRWRLLLLPPLPPLLLPAAEAAAASVSGAAAEKEDRCPVNGRNPCT
GQCPCAGWGCQCHGGRFLTGSSGFLTDGPNVYKTKCTWLJEGQPNRLMLR
FNHPATCSWHLVYDSDSLVAPLVAAPSLIIPERDGHETVEVATSGYALLHFF
SDAAYNLGNITNTFDMCPNCGRGECKISNSNTVQCESNWKGEACDIPHCVN
NGCFPHRGICNSDVRCSEFCSEVGPVFPANQSFWRTEYSDLKLPRASHKAV
VNGNIVVVGGMFNHSDYNMVLAYDLASHEWLNLSVNSVVRVYGHSLALYQDIY
MYGKIDSTGNTVVEVDLKNWTSILHLLPAAEQYAVVGHSAHIVLKNRVMWL
IFGHCPLYGYSVQVEVDLKNWTSILHLLPAAEQYAVVGHSAHIVLKNRVMWL
AFSANKRLADLVYDVTOMTILKDSRFFVLRVAVIVSGTMLVFGNTNDSM
SHGAKFSSDMAYDIDCDKSVLPRLDHDVNRFGHSAVLNRYNRYVFGNLSLL
SDILLVFSQCEAHQSEAAACLAAGPVRVCMWDTSSQCSWELAPEAKIEKSEPSK
RIPHRDCQNTCTANTNGWCNHCVPNTNHSCTEGQISIFKIDHCPKDNPMY
YCNKKTSCRSQALDQNCQWPRNOCIALPENGIGLHVLGNSCLITITAREYDNY
XLSCRNHNFLASITLTKQKVFVLRKLRIMQSSQMSKFTLPWVGLKINLVYQWE
DMSPFINSLOWMSEPSDAGFCGILSEPSTRGLKAATCINPLNGSCVCPANHSKQ
CRTPCALRTACGECTSGSSCEMCSNMKQCVDSNAYVAFPFQGCWENYTMSSCPEN

CSGYCTCSHLEQPCGCGWCTDPSNTGKCKIEGSKGPRVMPSPQPTGNSYVPQILLNS
SMCLDSRYNNSFHCPACQCNHSGKCIQINQICEKCNELITGKHCECTICISFYGDPTN
GKQPCPCNCHASLPCNTCKCTTKGKDEQCELVENYQGNPLTKGTCYATLL
IDYQFTLSLQDDRYXTAINFVATPDEQRDLDMFINASKFNELNLTWASFSAGTQ
AGEMPVVSXTNIIKEYKDSFSNEXDFRNHPNITFFVYVNGFTWPIKIQAFSOSHNF
MDLVQFFVTFSCFLSLLVAAVVKIKQSCWASRRERQLLREMOQWASPPFASVNA
LETDEPPDLGGSIKTVPKPIALEPCFGKNKAAVLVSFRLPRLGLGIPFGQSLAV
ASALVDLSQQMPVYKESGAVNRKQOPPAQPGTCT"

ORIGIN

Query Match 86.1%; Score 3694; DB 4; Length 4493;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 3947; Conservative 0; Mismatches 295; Indels 36; Gaps 4;

Qy 13 CGCGGGCAACTGAGGCAAGCTCAGGAGGAGGAGCGCGGCGAGCGCTGCGGGC 72

Db 58 CGCGGGCACTGAGGCGAGCT---GAGGAGGACCGTGGTGACACCGGCGCCACGAGC 114

Qy 73 AGGAGCGCGCGCGCGCACTGGGACCTGGGACCTGACACAGGGCTGGGAGCGCGGGCTCGGG 132

Db 115 AGGAACGCGCAGCGGCTGTCGGGACTGGGACCGGACTAGAGCTGGAAGCGCGGGCTCGCG 174

Qy 133 GCGGGCTGCGCTCCCGCGGCTGCTCTCCACCGCTGCGGCGCACCGCTGCTGCTG 192

Db 175 ACCGGCTGCGCTCCCGCGGCTGCTCTCCACCGCGG-----TGCGCGCGCTGCTG 213

Qy 193 CTGTTGTTGCTCCCGCGCGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252

Db 214 CTGCTGCTGCTCCGCGC-----GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264

Qy 253 GCGCGCGCGCGCGCGGCTGCTGCGGCTCAGCGCAGCGAGCGCAAGAAATGTACCGCGCC 312

Db 265 GCGGACGCGCGCTGCTGCGGCTCAGCGCAGCGAGCGCAAGAAATGTACCGCGCC 324

Qy 313 TGTCTAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372

Db 325 TGTGTCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384

Qy 373 GTGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTTAACCTGATCTTCTGCGTTT 432

Db 385 GTGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTTAACCTGATCTTCTGCAATT 444

Qy 433 GTGACAGATGCGCTCGAAATTAATAACAACAAGAGTGCAGCTGCTGCTGCTGCTGCTGCTG 492

Db 445 ATACAGATGCGCTCGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 504

Qy 493 CAGCCAAATAGAAATAGAGACTTCCTTTCAATTCATTTTGTACAGAGTGTAGTTGGGAC 552

Db 505 CAGCCAAATAGAAATAGAGACTTCCTTTCAATTCATTTTGTACAGAGTGTAGTTGGGAC 564

Qy 553 CATTTATATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGCTGCTGCTGCTG 612

Db 565 CATTTATATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGCTGCTGCTGCTG 624

Qy 613 CTCAATTTCTCTGAGAGAGATGCAATGAGACTTGCTCCCTGAGAGTGTGCTCCACATCAGT 672

Db 625 CTCAATTTCTCTGAGAGAGATGCAATGAGACTTGCTCCCTGAGAGTGTGCTCCACATCAGT 684

Qy 673 TATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732

Db 685 TATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744

Qy 733 TACAGTTTGTATGATGTCCTCAATTAACCTGCTCAGCGCGGAGAGTGTAGATTCAGTAAT 792

Db 745 TACAATTTGACATGCTGCTCAATTAATTTGCTCAGCGCGGAGAGATTAAGATCAGTAAT 804

Qy 793 AGACGCAAACTGTTGAATGTGAATGTTTCTGAAAACTGCAAAAGTGAAGCATGTGACATT 852

Db 805 AGCAGCAACACTGTTCAATGTGAATGTTTCTGAAAACTGCAAAAGTGAAGCATGTGACATT 864

Qy 853 CCTCACTGTACAGCAACTGTTGTTTCTCTCATGAGGACATCTGCAATTCAGTGTATGTC 912

Db 865 CCTCACTGTGTAATAACTGTTGTTTCTCTCATGAGGACATCTGCAATTCAGTGTATGTC 924

Qy 913 AGAGATGCTCCCTGCTTCTCAGACTGCGAGGCTCTCGGATGTCAGTTCCTGTCACAGCT 972
Db 925 AGAGGATGCTCCCTGCTTCTCAGAGTGGCAAGGTCCTCGATGTCAGTTCCTGTCACAGCT 984
Qy 973 AACCAAGTCAATTTGGGACTCGAGAGGAATATCTTAACCTTAAGCTCCCGAGAGCACTCAT 1032
Db 985 AACCAAGTCAATTTGGACAGAGAGGAATATCTCGACCTAAAGCTCCCGAGAGCTCTCAC 1044
Qy 1033 AAAGCTGTGTCAATGGAACCAATATATGCTGGTGTGGAGGATATATGTTCAACCACTCA 1092
Db 1045 AAAGCTGTGTCAATGGAACCAATATATGCTGGTGTGGAGGATATATGTTCAACCACTCA 1104
Qy 1093 GATTATAACATGCTTCTAGGATGACCTGCTTCTAGGAGTGGCTTCCCACTAAACCGT 1152
Db 1105 GATTATAACATGCTTCTAGGATGACCTGCTTCTAGGAGTGGCTTCCCACTAAACCGT 1164
Qy 1153 TCTGTGAACAATGCTGTTGTATGATATGCTCACTTCTTGGCAATATACAAGGATAAAT 1212
Db 1165 TCTGTGAACAATGCTGTTGTATGATATGCTCACTTCTTGGCAATATACAAGGATAAAT 1224
Qy 1213 TACATGTATGGGAGGAATATGATCAACTGGGATGTGACCAATGATGAGTGGAGTGT 1272
Db 1225 TACATGTATGGGAGGAATATGATCAACAGGGAATGTGACCAATGATGAGTGGAGTGT 1284
Qy 1273 CACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAGAGGAGCTATGCACTG 1332
Db 1285 CACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAGAGGAGCTATGCACTG 1344
Qy 1333 GTTGGGCACTCTGCACACATGTTTACACTGAGATGCGCGAGTGGTCATGCTGTCATC 1392
Db 1345 GTTGGGCACTCTGCAGACATGTTGACGTTGAAATATGCGCGAGTGGTCATGCTGTCATC 1404
Qy 1393 TTTGGTCACTGCGCTCTCTATGATATATGAGCAATGTGAGGATATGATTTGGATAG 1452
Db 1405 TTTGGTCACTGCGCTCTCTATGATATATGAGCAATGTGAGGATATGATTTGGATAG 1464
Qy 1453 AACACATGAGTATATACACACCGGCTGCGCTTGTGCAAGGAGTGGTACCGGCAATAGC 1512
Db 1465 AACACATGAGTATATACACACTGAGGCTGCGCTTGTGCAAGGAGTGGTACCGGCAATAGC 1524
Qy 1513 AGTGTGTTACGACATAGGACCGGCTTATAGTTCATGTTGCTGCTCAAGGCTTTCACT 1572
Db 1525 AGTGTCTATGATAGGACCGGCTTATAGTTCATGTTGCTGCTCAAGGCTTTCACT 1584
Qy 1573 GCCAATAGTACCGGCTTACAGATGATCTCTACCGATATGATGATGATACCGGATGAG 1632
Db 1585 GCCAATAGTACCGGCTTACAGATGATCTCTACAGATATGATGATGATACCGGATGAG 1644
Qy 1633 ACCATTTCTTAAGGACAGCGGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGGA 1692
Db 1645 ACCATTTCTTAAGGACAGCGGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGGA 1704
Qy 1693 ACCATGCTGTTTGGGGAACACACACATGACATCTATGAGGCTTGGCGCCAAA 1752
Db 1705 ACCATGCTGTTTGGGGAACACACACATGACATCTATGAGGCTTGGCGCCAAA 1764
Qy 1753 TGCTTCTCTTCAGATTTCAATGCTATGACATGCTGTGACCGGCTGCTGAGTGTCTCC 1812
Db 1765 TGCTTCTCTTCAGATTTCAATGCTATGACATGCTGTGACCGGCTGCTGAGTGTCTCC 1824
Qy 1813 AGACCTGATCTCCACATGATCAACAGATTTGGCCATTCAGCAGTCTTACACACAGC 1872
Db 1825 AGACCTGATCTCCACATGATCAACAGATTTGGCCATTCAGCAGTCTTACACACAGC 1884
Qy 1873 ACCATGATGTTGCTCGGTTGTTCAATAGTCTTCTCTCAGGACATCTCTGATTTACC 1932
Db 1885 ACCATGATGTTGCTCGGTTGTTCAATAGTCTTCTCTCAGGACATCTCTGATTTACC 1944
Qy 1933 TCGGAACAGTGTGATCGGATCGAGTGAAGCGGCTTGTGATGAGGAGGCTTCTTAAGGACCA 1992
Db 1945 TCGGAACAGTGTGATCGGATCGAGTGAAGCGGCTTGTGATGAGGAGGCTTCTTAAGGACCA 2004

Qy 1993 CGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGTTGGGCTGGCAACTGATGAA 2052
Db 2005 CGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGTTGGGCTGGCAACTGATGAA 2064
Qy 2053 CAAGAGAAAGTAAATCAGAAATGTTTCCAAAAGAACTCTTGACCATGACAGATGT 2112
Db 2065 C---AAGAAAGATAAATCAGAAATGTTTCTAAAGAAATTTTGACCATGACAGATGT 2121
Qy 2113 GACCACACACAGATTTGTTTACAGCTGTACAGCAACCAATGACTGCGCACTGTGCAAT 2172
Db 2122 GACCAGCTCACGATTTGTTTACAGCTGTACAGCAACCAATGACTGCGCACTGTGCAAT 2181
Qy 2173 GACCATGTTGTCCTCCAGGAAACACAGCTGTCTCAGAGGCGGAGATCTCCATTTTGTAT 2232
Db 2182 GACCATGTTGTCCTCCAGGAAACACAGCTGTCTCAGAGGCGGAGATCTCCATTTTGTAT 2241
Qy 2233 GAGAAATGTCCTCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAGCTGCGAGG 2292
Db 2242 GATCATTTGTCCTCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAGCTGCGAGG 2301
Qy 2293 TGTGCTCTCGACCAAGAACTGCGAGTGGGAGGCGGCGGATCAGGAGTGCATTCGCTGCC 2352
Db 2302 TGTGCTCTCGACCAAGAACTGCGAGTGGGAGGCGGCGGATCAGGAGTGCATTCGCTGCC 2361
Qy 2353 GAAATATCTGTGGCATTTGGCTGGCATTTGTTGGAACTCATGTTTGAATTTACTACT 2412
Db 2362 GAAATATCTGTGGCATTTGGCTGGCATTTGTTGGAACTCATGTTTGAATTTACTACT 2421
Qy 2413 GCCAAGGAAATATGACATGCTTAAATGTTCTGTAGAACCAATGCGCTTTGGCT 2472
Db 2422 GCCAAGGAAATATGACATGCTTAAATGTTCTGTAGAACCAATGCGCTTTGGCT 2481
Qy 2473 TCTCTTAAACCCAGAAAGGTAGAAATTTGCTTAAAGAGCTGCGAATTAATGCACTCA 2532
Db 2482 TCTCTTAAACCCAGAAAGGTAGAAATTTGCTTAAAGAGCTGCGAATTAATGCACTCA 2541
Qy 2533 TCTCAGAGCATCTCCAGCTCACCTTACCCCATCGGTCGCTTCCGAAGATCAATGTG 2592
Db 2542 TACAGGAGCATCTCCAGCTCACCTTACCCCATCGGTCGCTTCCGAAGATCAATGTG 2601
Qy 2593 TCTACTGTGTGCGGAAGATATGTCCTCCCATTTTACAAATAGTCTTACTACAGTGGATCGG 2652
Db 2602 TCTACTGTGTGCGGAAGATATGTCCTCCCATTTTACAAATAGTCTTACTACAGTGGATCGG 2661
Qy 2653 TCTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGAGCTG 2712
Db 2662 TCTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGAGCTG 2721
Qy 2713 AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCTTGCACCAAC 2772
Db 2722 AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCTTGCACCAAC 2781
Qy 2773 AGTGTAGCAGTGGCGGACCATGTCCTTGGAGGAGCATGTGGAGATTTGCACGAC 2832
Db 2782 AGTGTAGCAGTGGCGGACCATGTCCTTGGAGGAGCATGTGGAGATTTGCACGAC 2841
Qy 2833 GGCAGCTCTGAGTGTGCTGGTGGCAACCAATGAGCAGTGTGAGTCTCCAAATGCTCTAT 2892
Db 2842 GGCAGCTCTGAGTGTGCTGGTGGCAACCAATGAGCAGTGTGAGTCTCCAAATGCTCTAT 2901
Qy 2893 GTGGCTCTCTCTCTTGGCCAGTGTATGGAATGGAATAGTATAGTATGAGGACCTGCCCCCT 2952
Db 2902 GTGGCTCTCTCTCTTGGCCAGTGTATGGAATGGAATAGTATAGTATGAGGACCTGCCCCCT 2961
Qy 2953 GAAATTTGTTCAAGGCTTACTGTACTGTAGTCAATGCTTGGAGCAACCAAGGCTGTGGCTGG 3012
Db 2962 GAAATTTGTTCAAGGCTTACTGTACTGTAGTCAATGCTTGGAGCAACCAAGGCTGTGGCTGG 3021
Qy 3013 TGTACTGATCCAGCAATACTGGCAAGGAAATGTCATAGAGGTTCTATTAAGGACCA 3072
Db 3022 TGTACTGATCCAGCAATACTGGCAAGGAAATGTCATAGAGGTTCTATTAAGGACCA 3081
Qy 3073 GTGAAGATGCTTTCGCAAGGCTTACTAGGAAATTTCTATCCACAGGCCCCCTGCTCAATCC 3132

ITWATSFAGTQAGEBMPVSKTNIKEKYDSFNEKDFRHNPHNITPFFVVSNETWBI
KQIAPFSORNSMDLVQFVETPFCFLSILLVAAVWKIKOSWASRREROLLEMOO
MASRPPASVVALETTDEBPPDLIGSITKVPKPALEPCFNKAAVLSVFLPRGLG
GIPIPGQSLAVASALVLSQMPVIYKEKSGAVNRKQPPAQPGICI"

ORIGIN

Query Match		83.2%;	Score 3570.4;	DB 10;	Length 4284;
Best Local Similarity		89.7%;	Pred. No. 0;		
Matches 3848;		Conservative 0;	Mismatches 436;	Indels 6;	Gaps 1;
QY	1	ATGCTGGCGCGAGCGCGCACTGAGGCAAGGCTGAGGAGGAGGAGCGCGCGCGCA	60		
DB	1	ATGCTGGCGCGCGCGCGCACTGAGGCGCTGCTGAGGGGAGCACCGCGCGCAGCT	60		
QY	61	GCCTCGCGCGCAGGAGCGCGCGCGCACTGCGGACTGCGGAGCGCTGGGAGG	120		
DB	61	GCSCCGCGCGCGAGGAGGCTGAGCAGTACCGACTCTGCGCGCGCGCTGGGAGG	120		
QY	121	CCGGGCGCGCGCGCGCGCTGCGCTCCCGCGCGCTGCTCCACCGCTGCGCGCGG	180		
DB	121	CCGGGCGCGCGCGCGGATTTGTCTCCCGTGGGTGTGCGGGCTCTGCCCGCGCG	180		
QY	181	CTGCTGCTGCTGTTGTGCTCCCGCGCGCTGTTGCTGCTGCTGCTGCTGCTG	240		
DB	181	CTGCTGCTGCTGCTGCTCT-----TTGCTACTGCTGCTGCTGCTGCTGCTG	234		
QY	241	GCCGAGCGCGCGCGCGCGCGGCTGCGGCTCAGCGCAGCGCGCGCGCAAGAA	300		
DB	235	GCCGAGCGCGCGCGGCTGCGCGCGGCTGAGCTGCGCGCGCGCGCGCGCAAGAA	294		
QY	301	TGTGACCGCGCTGTGTCAACCGCGCTGCTGCTCAACCGCTGCGCGCGCGCTG	360		
DB	295	TGTGACCGCGCTGTGTCAACCGCGCGCTGCTCAACCGCTGCGCGCGCTGCTG	354		
QY	361	CCGCGCGCTGGTGGCGAGCAATGCGAGCTGCGGCGCGCGCTTCCAGTACTGGA	420		
DB	355	CCGCGCGCTGGTGGCGAGCAATGCGAGCTGCGGCGCGCGCTTCCAGTACTGGA	414		
QY	421	TCCTCTGGGCTTGTGACAGTGCACCTGGAATTTAAATAAACAAGAGTGCACCTG	480		
DB	415	TCCTCTGGATTCGTAAACAGTGCACCTGGGAATTTAAATTAACAAGAGTGCAC	474		
QY	481	CTCATTTAGGAGCAGCAATAGATTAATGAGACTTCTGTTCAATCAATTTGCTAC	540		
DB	475	CTCATTTAGGAGCAGCAATAGATTAATGAGACTTCTGTTCAATCAATTTGCTAC	534		
QY	541	TGTAGTTGGGACCAATTTATGTTATGATGGGACTCAATTTATGCAACCGCTAG	600		
DB	535	TGTAGTTGGGACCAATTTATGTTATGATGGGACTCAATTTATGCAACCGCTAG	594		
QY	601	GCATTTAGTGGCTCATTTGCTGAGAGAGATGGCAATGAGCTGCTCCCTGAGG	660		
DB	595	GCATTTAGTGGCTCATTTGCTGAGAGAGATGGCAATGAGCTGCTCCCTGAGG	654		
QY	661	GCCACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTTAAATTTGACTGGA	720		
DB	655	GTCATTTAGTATGCACTGCTGCAATTTTTCAGTATGCTGCTTCAATCTGACTG	714		
QY	721	TTTAAATTAATTAAGTTTGAATGCTGCAATTAATGCTGAGCGCGAGGAGTGT	780		
DB	715	TTTAAATTAATTAAGTTTGAATGCTGCAATTAATGCTGAGCGCGAGGAGTGT	774		
QY	781	AAGATCAGTAAATAGCAGCAAACTGTTGAATGTAATGTTCTGAAATCTGAAAG	840		
DB	775	AAGAGCAGTAAACAGCAGCAAACTGTTGAATGTAATGTTCTGAAATCTGAAAG	834		
QY	841	GCATTTGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900		
DB	835	TCGCTGTGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	894		
QY	901	TCAAGTGTGCTGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960		
DB	895	GCAAGTGTGCTGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	954		

QY	961	CTGTACAGCTAAACAGCTCAATTTGGACTCGAGAGGAAATTTCTAACTTAAAGCTCCC	1020		
DB	955	CTGTACAGCTAAACAGCTCAATTTGGACTCGAGAGGAAATTTCTAACTTAAAGCTCCC	1014		
QY	1021	AGAGCATCTCAATAAGCTGTGTCAATGGAACAATTAATGTGGGTGTGGAGGATATATG	1080		
DB	1015	AGAGCATCTCAATAAGCTGTGTCAATGGAACAATTAATGTGGGTGTGGAGGATATATG	1074		
QY	1081	TTCAACCACTCAGATTATAAATGCTTCTAGCGTATGACCTTCTTCTAGGAGTGGCTT	1140		
DB	1075	TTCAACCACTCAGATTATAAATGCTTCTAGCGTATGACCTTCTTCTAGGAGTGGCTT	1134		
QY	1141	CCACTAAACCGTCTGTGAACAATGTGTGTGATATGCTCATTTCTTGGCAATATAC	1200		
DB	1135	CCCTAAACCACTCTGTGAACAATGTGTGTGATATGCTCATTTCTTGGCAATATAC	1194		
QY	1201	AAGGATAAAATTTACATGTATGGAGGAAATTTGATCCAACTGGGAATGTCAACAATGAG	1260		
DB	1195	AAGGATAAAATTTACATGTATGGAGGAAATTTGATTTCAACAGGAAATGTCAACAATGAG	1254		
QY	1261	TTGAGAGTTTTCACATTTCAATGAGTCATGGGTGTGTTGACCCCTTAAGGCAAGAG	1320		
DB	1255	CTGAGAGTATTTCAATTTCAATGAGTCATGGGTGTGTTGATCCAAACCAAGCAAGAG	1314		
QY	1321	CAGTATGAGTGTGGGCACTCTGCACACATTTGACACTGAAGAAATGGCCGAGTGTCTC	1380		
DB	1315	CAGTATGAGTGTGGGCACTCTGCACACATTTGACACTGAAGAAATGGCCGAGTGTCTC	1374		
QY	1381	ATGCTGGTCACTTTTGGTCACTGCGCTCTCTATGAGTATATAAGCAATGTGACAGAAAT	1440		
DB	1375	ATGCTGGTCACTTTTGGTCACTGCGCTCTCTATGAGTATATAAGCAATGTGACAGAAAT	1434		
QY	1441	GATTTGGATAAACAACATGGAGTATTTACACACCAAGGTCCTTGTGCAAGGGGT	1500		
DB	1435	GATTTGGATAAACAACATGGAGTATTTACACACCAAGGTCCTTGTGCAAGGGGT	1494		
QY	1501	TAGGCGCATPAGCAGTGTTTTACGACCAATAGGACCAAGGCGCTTATAGTTTATGGTGTAC	1560		
DB	1495	TATGGCGCATPAGCAGTGTTTTACGACCAATAGGACCAAGGCGCTTATAGTTTATGGTGTAC	1554		
QY	1561	AAGGCTTTCAGTGGCAATAAGTACCGCTTGCAGATGATCTACCGATATGATGTGAT	1620		
DB	1555	AAGGCTTTCAGTGGCAATAAGTACCGCTTGCAGATGATCTACCGATATGATGTGAT	1614		
QY	1621	ACCGAGTGTGACCACTTTTAAAGACAGCGGATTTTCCGTTACTTTGACACACAGCTGTG	1680		
DB	1615	ACCGAGTGTGACCACTTTTAAAGACAGCGGATTTTCCGTTACTTTGACACACAGCTGTG	1674		
QY	1681	ATAGTGTGAGGACCACTGCTGTGTTGGGGGAAACACACACATGACATCTATGAGC	1740		
DB	1675	ATAGTGTGAGGACCACTGCTGTGTTGGGGGAAACACACACATGACATCTATGAGC	1734		
QY	1741	CATGGCGCAAAATGCTTCTTTCAGATTTTATGGCCTTATGACATTTGCTGTGACCGCTGG	1800		
DB	1735	CATGGCGCAAAATGCTTCTTTCAGATTTTATGGCCTTATGACATTTGCTGTGACCGCTGG	1794		
QY	1801	TGAGTGTTCACAGCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC	1860		
DB	1795	TGAGTGTTCACAGCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC	1854		
QY	1861	TTACACAAACAGCACTGATGTGTGCTGGTGTGTTCAATAGTCTCTCTCAGCGGACATC	1920		
DB	1855	TTGCAACAACAGTACATGATGTGTGCTGGGCTTTCAACAGCTCTCTCAGTGAAGTC	1914		
QY	1921	CTGGTATTTACCTCGGACAGTGTGATCGGATCGGATCGGCTGTGTTTACGAGCA	1980		
DB	1915	TTAGTCTTTTACCTCGGACAGTGTGATCGGATCGGCTGTGTTTACGAGCA	1974		
QY	1981	GGACCTGTGTATTCGGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCTGGCGCTG	2040		
DB	1975	GGACCTGTGTATTCGGTGTGTGGAACACAGTGTGTCTCAGTGTATCTCTGGCGCTG	2034		

QY 2041 GCAACTGATGAAACAAGAGAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
DB |||||
DB 2035 GCAACTGAGAAACAAGAGAGAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2094
QY 2101 CATGACAGATGTGACACAGACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGC 2160
DB |||||
DB 2095 CATGACAGATGTGACACAGACACAGATTTGTTACAGCTGTACAGCCAAATACCAATGACTGC 2154
QY 2161 CACTGAGTGAATGACCAATTTGTTGCCAGAGAACACAGCTGTGACAGAGCCAGATCTCC 2220
DB |||||
DB 2155 CACTGAGTGAATGATCACTGTGTCTCCGTGTAACCAAGCTGTGACAGAGCCAGATCTCC 2214
QY 2221 ATTTTATGATGATGAGAAATTTGCCCAAGGATAACCCATGCTACTGTAAACAGAGACC 2280
DB |||||
DB 2215 ATTTCAAGTATGACAAATTTGCCCAAGGATAACCCATGCTACTGTGCAATTAAGAAACA 2274
QY 2281 AGTGCAGAGGTGTGCCCTGGAACCAAGAACTGCCAGTGGAGCCCGGAATCAGGAGTGC 2340
DB |||||
DB 2275 AGTGCAGAGGTGTGCCCTTAGACCAAGAACTGCCAGTGGAGCCCGGAATCAGGAGTGC 2334
QY 2341 ATTGCCCTGCCGAAATATCTGTGCCATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2400
DB |||||
DB 2335 ATGCTCTGCCCTGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACTCATGTCTG 2394
QY 2401 AAAATTAATCTGACCAAGGAGAAATATGACAAATGCTAAATTTGTTGTAGGAACCACAAT 2460
DB |||||
DB 2395 AAAATCACTACTGCTAAGGAAATTTATGACAAATGCTAAATTTGTTGTAGGAACCACAAT 2454
QY 2461 GGCCTTTTGGCTTCTTTACACCCAGAGAGAGGTAGAAATTTGCTTAAAGCAGCTGGGA 2520
DB |||||
DB 2455 GGCCTTTTGGCTTCTTTACATCCAGAGAGAGGTGGAAATTCGTCTTAAAGCATCTTGA 2514
QY 2521 ATAAATGCACTCATCTCAGAGCATGTCCAAGTCTCACTTTAAACCCATGCTGGCTTCGG 2580
DB |||||
DB 2515 TTAATGCACTCATCTCAGAGCATGTCCAAGTCTCACTGTCTCCTCAGAGCTTCGG 2574
QY 2581 AAGATCAATGTGCTCTTATTTGCTGGGAGGATATGCTCCATTTACAAATAGTTTCTG 2634
DB |||||
QY 2641 CAGTGGATGCTGTGACCCAGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2700
DB |||||
DB 2635 CAGTGGATGCTGTGACCCAGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2694
QY 2701 ACTGGGATGCTGTGACCCAGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2760
DB |||||
DB 2695 ACTGGGATGCTGTGACCCAGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2754
QY 2761 CAGTGGATGCTGTGACCCAGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2820
DB |||||
DB 2755 CAGTGGATGCTGTGACCCAGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2814
QY 2821 GATGGACAGGCGGAGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 2880
DB |||||
DB 2815 GATGGACAGGCGGAGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 2874
QY 2881 TCCATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 2940
DB |||||
DB 2875 TCCATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 2934
QY 2941 ACCTGCCCCCTGAAATTTGCTGAGGCTACTGTACCTGTGATGCTGTGATGCTGTGATGCTG 3000
DB |||||
DB 2935 AGCTGCCCCCTGAAATTTGCTGAGGCTACTGTACCTGTGATGCTGTGATGCTGTGATGCTG 2994
QY 3001 GCTGTGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 3060
DB |||||
DB 2995 GCTGTGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 3054
QY 3061 TATAAGAGCCAGTGAAGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTG 3120
DB |||||
DB 3055 TATAAGAGCCAGTGAAGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTG 3114
QY 3121 CTGCTCAATTCAGCATGTGTGTAGAGGACAGAGATACAACTGGTCTTTCTTCACTCTGT 3180

DB 3115 CTTCTGAATCTAGCATGTGTTTAGAGGACAGAGATACAACTGGTCTTTCACTCTGT 3174
QY 3181 CCAGCTTGGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGT 3240
DB |||||
DB 3175 CCAGCTTGGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGT 3234
QY 3241 GAGAACTGACCCACAGGCAAGCACTGCGAGACCTGCGATATCTGGCTTCTACGGTGATCCC 3300
DB |||||
DB 3235 GAGAACTGACCCACAGGCAAGCACTGCGAGACCTGCGATATCTGGCTTCTACGGTGATCCG 3294
QY 3301 ACCAATGAGGGAATGTCAGCCATGCAATGCAATGCGGCAACGCTCTCTGTGCAACAC 3360
DB |||||
DB 3295 ACTAATGAGGGAATGTCAGCCATGTAAGTGCAACGGGCAATCATCGCTGTGCAATAC 3354
QY 3361 AACACGGGCAAGTGTCTGACACCAACGAGGGGCTCAAGGGGACAGTGCCAGCTATGT 3420
DB |||||
DB 3355 AACACGGGCAAGTGTCTGACACCAACGAGGGTGTCAAGGAGATGAGTGCCAGCTATGT 3414
QY 3421 GAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTTATATCTTTCT 3480
DB |||||
DB 3415 GAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTTATATACCTTCTC 3474
QY 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAGATGATCGCTATTACAGCATTC 3540
DB |||||
DB 3475 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAGATGATCGCTATTACAGCATTC 3534
QY 3541 AATTTTGTGGTACTCTGACGAAACAAACAGGGATTTGGAATGTTCAATCAATGCTCTC 3600
DB |||||
DB 3535 AATTTTGTGGTACTCTGATGAAACAAACAGGGATTTGGAATGTTCAATCAATGCTCTC 3594
QY 3601 AAGAAATTTCAACCTCAACATCCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGA 3660
DB |||||
DB 3595 AAAAACTTCAACCTCAACATCCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGA 3654
QY 3661 GAGAGATGCTGTGTTTCAAAAAACCAACATTTAAGGAGTACAAAGATGTTTCTCTAAT 3720
DB |||||
DB 3655 GAGAAATGCTGTGTTTCAAAAAACCAACATTTAAGGAGTACAAAGATGTTTCTCTAAT 3714
QY 3721 GAGAAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAAATTC 3780
DB |||||
DB 3715 GAGAAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAAATTC 3774
QY 3781 ACTGCCCCATCAAAATTCAGATGCTCTCTCAGCAGCAGCAATTTTATGAGCTGTGA 3840
DB |||||
DB 3775 ACTGCCCCATCAAAATTCAGATGCTCTCTCAGCAGCAGCAATTTTATGAGCTGTGA 3834
QY 3841 CAGTTCTTCTGTCATTTCTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
DB |||||
DB 3835 CAGTTCTTCTGTCATTTCTTCTCAGCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3894
QY 3901 TCGAAGATCAAAACAAAGTGTGTGGGCTCCAGACCTAGAGAGCACTTTCTTGGAGATG 3960
DB |||||
DB 3895 TCGAAGATCAAAACAAAGTGTGTGGGCTCCAGAGGAGGAGCACTTTCTTGGAGATG 3954
QY 3961 CAACAGATGGCCAGCGCTCCCTTTGCTCTCTGTAATGTCGCTTTGAAACAGATGAGAG 4020
DB |||||
DB 3955 CAGCAGATGGCCAGCGCTCCCTTTGCTCTCTGTAATGTTGCTTTGAAACAGATGAGAG 4014
QY 4021 CTTCTGATCTTTTGGGGGAGTATAAGAGTGTTCCTTCCAAACCAATGTCAGCTGGAGCCG 4080
DB |||||
DB 4015 CTTCTGATCTTTTGGGGGAGTATAAGAGTGTTCCTTCCCAAGCCCAATTTGCTCTGGAGCC 4074
QY 4081 TTTTGGCAACAAAGCCGCTGCTCTCTCTGTTTGTAGGCTCCCTCGAGGCTCTGGGT 4140
DB |||||
DB 4075 TGTCTTGTAAACAAAGCTGAGTCTCTCTGTAATTTGTGAGGCTCCCTCGAGGACTGGGT 4134
QY 4141 GGCATCCCTCTCTCTGGGAGTCAAGTCTTCTGTGCGCAGGCTCTGTGCGAATTTCT 4200
DB |||||
DB 4135 GGAATCCCTCTCTCTGGTCAAGTCTTCTGTGCGCAGTGTGCTGTGCGAATTTCT 4194
QY 4201 CAGCAGATGGCGATGTGTACAGGAGAGTCAAGGCGCTGTGAGAAACCCGAGACAGAG 4260

QY	1023	AGCATCTCATAAAGCTGTGGTCAAATGGAAACATTAATGTGGGTGTGTGGAGGATATATGTT	1082
Db	1089	AGCCTCTCATAAAGCTGTGGTCAAATGGAAATATAATATGGGGTGTGTGGCGGATATATGTT	1148
QY	1083	CAACCACTCAGATTATAACATGGTTCTAGCGGTATGACCTTGCTTCTTAGGGAGTGGCTTCC	1142
Db	1149	CAACCACTCAGATTACAGCATGGTTTAAAGCTATGACCTGACTTCTAGGGAATGGCTTCC	1208
QY	1143	ACTAAACCGTCTGTGGAACAATGTGGTGTTAGATATAGTCAATCTCTTTGGCATATACAA	1202
Db	1209	ACTAAACCAATCTGTGAAACAGTGTGGTGTGTAGATATAGTCAATCTTTGGCATATACATAA	1268
QY	1203	GGATAAAATTTACATGTATGAGGAGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT	1262
Db	1269	GGATAAAATCTACATGTATGAGGAGAAAAATTTGATTTCAACAGGGAACGTGACCAATGAGCT	1328
QY	1263	GAGAGTTTTTCACATTCATATGAGTCATGGGTGTGTGACCCCTTAAGSCAAAGAGCA	1322
Db	1329	GAGAGTATTTTCATATTTCAATAATGAATCATGGGTATGTTAACTCGGAAGCTAAGGATCA	1388
QY	1323	GTATCCAGTGGTGGCCACTCTGCACACATTTGTTTACACTGAAGAATGCGCGAGTGGTCAT	1382
Db	1389	GTATCCAGTGGTGGACACTCAGCACACATTTGTTTACACTGCGCATCTGGCCGTGTGGTCAT	1448
QY	1383	GCTGTGTATCTTTGGTGTCTAGTCCCTCTCTATGAGATATAAAGCAATGTGCAGGAATATGA	1442
Db	1449	GTGTGTCTATCTTCGGTCTATTCGCCACTCTATGATATATAAGGGTTGTGCAGGAATATGA	1508
QY	1443	TTTGGATAAAGAACACATGGAGTATATTACACACCAGGGTGCCTGTGTGCAAGGGGGTTA	1502
Db	1509	CTTGGAAAAGAACACATGGAGTATATTACATCTCAGGGTGCTCTTGTGTCAAGGGGGTTA	1568
QY	1503	CGGCCATAGCAGTGTTTACGACCATAGGACCAAGGCCCTTATACGTTTCATGGTGGCTACAA	1562
Db	1569	TGGCCACAGTAGTGTTTATGATGACAGGACCAAGGGCTCTGACGTTTCATGGTGGCTACAA	1628
QY	1563	GGCTTTTCAGTCCCAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC	1622
Db	1629	GGCTTTTCAGCGCCAAACAAATACCGGCTTGAGATGACCTCTACAGATACGATGTGGATAC	1688
QY	1623	CCAGATGTGGACCATCTTTAAGGACAGCCGATTTTCCGTTACTTGCACACAGCTGTGAT	1682
Db	1689	TCAGATGTGGACCATCTTTAAGGACAGCCGATTTTCCGTTACTTGCATACAGCTGTGAT	1748
QY	1683	AGTCAGTGGACCAATGCTGTGTGTGGGGAAACACACAAATGACACATCTATGAGCCA	1742
Db	1749	AGTCAGTGGAAACCAATGCTGTGTGTGGGGAAACACACAAATGACACTTCCATGAGCCA	1808
QY	1743	TGGCGCCAAATGCTTCTTTTCAGATTTTCATGGGCTATGACATTCCTGTGACCGTGGTTC	1802
Db	1809	CGGTGCCAAATGCTTCTCTCCGACTTTCATGGCTTATGACATTTGTTGACCGATGGTTC	1868
QY	1803	AGTGCCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGTCTTT	1862
Db	1869	AGTGCCTTCCAGACCTGAGTCTCATCATGATGTCAAACAGATTTGGCCATTCAGAGTCTTT	1928
QY	1863	ACACACAGCACCATGATGTGTTCGGTGTGTTTCAATAGTCTCTCTCCTCAGGCGCATGCT	1922
Db	1929	GTACACAGCACCATGATGTGTGTTCGGCGGCTTCAACAGCCTCTCTCTCAGTGAAGTCTTT	1988
QY	1923	GGTATTCACCTCGGAACAGTGTGATCGGNTCGGAGTGAAGCGGCTTGTTTAGCAGCAGG	1982
Db	1989	GGTCTTTTACCTCGGAGCAGTGGCATGCACCCGCACTGAAGTGTGTGTGGCAGCAGG	2048
QY	1983	ACCTGGTATTCGGTGTGTGTGGAACACAGGGTCTCTCAGTGTATCTGTGGCGCTGGC	2042
Db	2049	ACCTGGTATTCGGTGTGTGTGGGACACACAGTCTGATCTCGATGTAATCTCTGGAGTGTGGC	2108
QY	2043	AATCTGATGACACAGAGAAAAGTTAAATTCAGAAATGTTTTTCCAAAAGAACTCTTTGACCA	2102
Db	2109	AATCTGAAGAACACAGCAGAAAAGTTAAATTCAGAGTGTTTTTCTAAAAGAACCTTTGACCA	2168

QY	2103	TGACAGATGTGACCAAGCACA	CAGATTTGTTACAGCTGTACACCAACCAATGACTGCCA	2167
DB	2169	TGACAGATGTGACCAAGCACA	CAGATTTGTTACAGCTGTACACCAACCAATGACTGCCA	2228
QY	2163	CTGGTGCATGACCAATCTGTGTC	CCCAAGAACCAACAGCTGTCTAGAAGCCAGATCTCCAT	2222
DB	2229	CTGGTGCATGATCTACTGTGTC	CCCTGTGGAACACAGCTGTACAGAAGCCAGATCTCCAT	2288
QY	2223	TTTTAGGTATGAGAAATGCC	CCCAAGGATAACCCCATGTACTCTGTAAAGAAAGCCAG	2282
DB	2289	TGCCAAGTATGAGAGTTGCC	CCCAAGGATAACCCCATGTACTCTGTAAAGAAAGCCAG	2348
QY	2283	CTGCAGGAGCTGTGCCCTG	AGCAGAACTGCCAGTGGAGCCCGGAATCAGAGATGCAAT	2342
DB	2349	CTGCAGGAGCTGTGCCCTG	AGCAGAACTGCCAGTGGAGCCCGGAATCAAGAGTGCAT	2408
QY	2343	TGCCCTGCCCAAAATATCTGT	GGCAATTTGGTTGGAAACTCATGTTTGAA	2402
DB	2409	CGCCCTGCCGAAATATCTGT	GCAATGGCTGGCAATTTGGTTGGAAACTCTGTGTCTGAA	2468
QY	2403	AATTACTACTCCAAAGGAGAT	TAATGACAATTGCTTAATTTCTTGTAGGAACCAACAATGC	2462
DB	2469	AATCACTACTCTAAGGAGAA	TTATGACAATTTGCTCTGTAGGAACCAACAATGC	2528
QY	2463	CTTTTGGCTTCTCTTCAAC	CCCAAGAAAGTAGAATTTGCTCTTAAGCAGCTGCGAAT	2522
DB	2529	CTTTTGGCTTCTCTCATCT	CCCAAGAAAGTGGAGTTTGTCCTTAGCAGCTTCGATT	2588
QY	2523	AATGAGTCTATCTCAGAGCAT	GTCCAAGTCACTTAACCCCATGGCTCGGCCCTTCGGAA	2582
DB	2589	AATGCAATCATCTCAAAAG	TATGTCCAAGTCACTCTGACTCCATGGGTTGGTCTTCGGAA	2648
QY	2583	GATCAATGTGTCTACTGTG	CTGGGAAGATATGTCCTCCATTTTACAATAGTTTACTACA	2642
DB	2649	GATCAATGTGTCTACTGTG	CTGGGAAGATATGTCCTCAATCACAATAAGTTTGTCTGCA	2708
QY	2643	GTGATGCCCTCTGAGCCCA	GATGATGCGAATTTCTGTGGAAATTTTACAGAACCCAGTAC	2702
DB	2709	GTGATGCCATCTGAGCCCA	GATGATGCGCTTCTGTGGATCTTGTAGAGCCTAGTAC	2768
QY	2703	TCGGGACTCAAGGCTGCAC	CTGCTCAACCCATCAATGTAGTGTCTGTGAAAGGCC	2762
DB	2769	TCGGGGATTAAGGCTGCAC	CTGCTCAACCCCTCTCAATGGCAGCGTCTGTGAAAGGCC	2828
QY	2763	TGCAAAACCAAGTGTAAAG	CAGTGCCTGGACAACCATGTGCTTTGAGACACGATGTGGAGA	2822
DB	2829	TGCAAAACCAAGTGCACAG	CACTGCGCTTCTGGGACACATGTGCTTGGGACAGCTGTGGCGA	2888
QY	2823	TTGCACAGCGGAGCTCTG	AGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAATC	2882
DB	2889	GTGCATCTAGCAGAGCTCG	AGTGCATGTGGTGCAGTAACTGAAGCAGTGTGTGGAATC	2948
QY	2883	CAATGCTATGTGGCTCTCT	CTCCCTTTGCGCAGTGTATGGAAATGGTATACCATCAGCAC	2942
DB	2949	CAATGCTACTGTGGCTCT	CTCTTTTGGCAGTGTATGGAAATGGTATACCATGAGCAG	3008
QY	2943	CTGCCCCCTTGAAAAATT	GTTCAGGCTACTGTACTGTAGTCAATCTGTTGGAGCAACCCAGG	3002
DB	3009	CTGCCCCCTTGAAAAATT	GTCTCTGGCTACTGTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGG	3068
QY	3003	CTGTGCTGTGTACTGATCC	AGCAATACTGGCAAGGGAATGCATAGAGGGTTCCTTA	3062
DB	3069	CTGTGTGTGTGTACTGAT	CTCTAGCAATACTGGGAAGGAAATGTATTAGAGGGCAGCTA	3128
QY	3063	TAAAGGACCAAGTGAAGAT	GCCTTCGCAAGCCCTCAGGAAATTTCTATCCACAGCCCCCT	3122
DB	3129	TAAAGGACCTGTGAAGAT	GCCTTCAGGCTCTGCAGGAAATGTATCCACAGCCCCCT	3188
QY	3123	GCTCAATTCAGCATGTGTCT	AGAGCAGCAGATACAACTGGTCTTTTCAATTCATCTGTCC	3182
DB	3189	TCGAAACTTCAGCATGTGT	CTAGAGCAGCAGATACAACTGGTCTTTTCAATTCATCTGTCC	3248
QY	3183	AGCTTGCCAAATCAACGG	CCACAGTAAATGTACATTCAGAGCATCTGTGAGAAAGTGTGA	3242

2163 QY CTGGTGCATGACCAATTGTGTCCCAGGAACCAAGCTGCTCAGAAGGCCAGATCTCCAT 2222
2232 Db CTGGTGCATGATCACTGTGTCCCTGTGAACCAAGCTGCACAGAAGGCCAGATCTCCAT 2291
2223 QY TTTTAGGTATGAGAAATGCCCCCAAGGATAACCCCATGTACTACTGTGTACACAGAGACCCAG 2282
2292 Db TGCCAAATGATGCAATGCCCCCAAGGATAACCCCATGTACTACTGTGTACACAGAGACCCAG 2351
2283 QY CTGCAAGAGCTGTGCTGCTGGACCAAGCTGCAAGTGGAGGCCGCCGGAATCAGGAGTGCAT 2342
2352 Db CTGCAAGAGCTGTGCTGCTGGACCAAGCTGCAAGTGGAGGCCGCCGGAATCAGGAGTGCAT 2411
2343 QY TGCCCTGCCCCGAATAATCTGTGGCAATTTGGCTGGCAATTTGGTGTGGAATCACTCATGTTTGA 2402
2412 Db CGCCCTGCCCCGAATAATCTGTGGCAATTTGGCTGGCAATTTGGTGTGGAATCACTCATGTTTGA 2471
2403 QY TAATTAATCTGCTCAAGGAGAAATTAAGCAATGCTAAATTTGTTCTGTAGGAACCAACATGC 2462
2472 Db TAATTAATCTGCTCAAGGAGAAATTAAGCAATGCTAAATTTGTTCTGTAGGAACCAACATGC 2531
2463 QY CTTTTTGGCTCTCTTTACCAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGCTGCGAAT 2522
2532 Db CTTTTTGGCTCTCTTTACCAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGCTGCGAAT 2591
2523 QY AATGAGTCACTCAGAGCATGTCCAGCTCACTTAACCCCATGGTTCGGCTTCGGAA 2582
2592 Db AATGAGTCACTCAGAGCATGTCCAGCTCACTTAACCCCATGGTTCGGCTTCGGAA 2651
2583 QY GATCAATGTGCTCTACTGCTGGGAGAGATATGTCCTCATTTTACAAATAGTTTACTACA 2642
2652 Db GATCAATGTGCTCTACTGCTGGGAGAGATATGTCCTCATTTTACAAATAGTTTACTACA 2711
2643 QY GTGGATGGCTGTGAGCCAGTGTGATGCTGGATTTCTGTGGAATTTATCAGAACCCAGTAC 2702
2712 Db GTGGATGGCTGTGAGCCAGTGTGATGCTGGATTTCTGTGGAATTTATCAGAACCCAGTAC 2771
2703 QY TCGGGGACTGGAAGTGTCAACTGTCATCAACCTCATCAATGTAGTCTGTGAAAGGCC 2762
2772 Db TCGGGGACTGGAAGTGTCAACTGTCATCAACCTCATCAATGTAGTCTGTGAAAGGCC 2831
2763 QY TGCAAAACCAAGTGTAAAGCAGTGCAGGACACCATGTGCTTTGAGGACAGCATGTGGAGA 2822
2832 Db TGCAAAACCAAGTGTAAAGCAGTGCAGGACACCATGTGCTTTGAGGACAGCATGTGGAGA 2891
2823 QY TTGCAACCCAGCAGCTCTGAGTGCAATGTGGTGACACACATGAGCAGTGTGGACTC 2882
2892 Db TTGCAACCCAGCAGCTCTGAGTGCAATGTGGTGACACACATGAGCAGTGTGGACTC 2951
2883 QY CAATGCTATGTGGCT 2942
2952 Db CAATGCTATGTGGCT 3011
2943 QY CTGCCCCCTGAAATTTGTTCAGGCTATGTAACCTGTAGTCAATGCTTTGGAGCAACCCAGG 3002
3012 Db CTGCCCCCTGAAATTTGTTCAGGCTATGTAACCTGTAGTCAATGCTTTGGAGCAACCCAGG 3071
3003 QY CTGTGCTGCTGTACTGTATCCAGCAATCTGCGCAAGGGAATGCAATAGAGGGTTCTTA 3062
3072 Db CTGTGCTGCTGTACTGTATCCAGCAATCTGCGCAAGGGAATGCAATAGAGGGAGCTA 3131
3063 QY TAAAGGACCAAGTGAAGATGCTTCCAGAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122
3132 Db TAAAGGACCAAGTGAAGATGCTTCCAGAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3191
3123 QY GCTCAATTCAGCATGTGCTAGAGACAGCATACACTGCTTTCTTCAATTCATCTGTC 3182
3192 Db GCTCAATTCAGCATGTGCTAGAGACAGCATACACTGCTTTCTTCAATTCATCTGTC 3251
3183 QY AGCTTGCCCAATGCAACCGGCCAGTAAATSCATCAATCAGACATCTGTGAGAGTGTGA 3242
3252 Db AGCTTGCCCAATGCAACCGGCCAGTAAATSCATCAATCAGACATCTGTGAGAGTGTGA 3311

3243 QY GAACCTGACCAAGCAAGCACTGCGAGACCTGCTCATATCTGGCTTCTACGGTATCCAC 3302
3312 Db GGACCTGACCAAGCAAGCACTGCGAGACCTGCTCATATCTGGCTTCTATGGTATCCAC 3371
3303 QY CAATGAGGAAATGTGAGCAATGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCA 3362
3372 Db TAATGAGGAAATGTGAGCAATGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCA 3431
3363 QY CACGGGCAAGTGTCTTGCAACCAAGGGGCTCAAGGGGAGAGTGGCCAGCTATGTGA 3422
3432 Db CACTGGCAAGTGTCTTGCAACCAAGGGGCTCAAGGGGAGAGTGGCCAGCTATGTGA 3491
3423 QY GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATTAATCTCTCTTAT 3482
3492 Db GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATTAATCTCTCTTAT 3551
3483 QY TGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGCTGCTATTAACAGCTATCAA 3542
3552 Db TGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGCTGCTATTAACAGCTATCAA 3611
3543 QY TTTTGGCTACTCTCTGACGAACAAACAGGGATTTGGACATGTTTCAATGCTCTCAA 3602
3612 Db TTTTGGCTACTCTCTGACGAACAAACAGGGATTTGGACATGTTTCAATGCTCTCAA 3671
3603 QY GAATTTCAACCTCAAATCAACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3662
3672 Db GAATTTCAACCTCAAATCAACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3731
3663 QY AGAGATGCTGCTGCTTCAAAAAACCAATTAAGGATGACAAAGATAGTCTTCTTAATGA 3722
3732 Db AGAGATGCTGCTGCTTCAAAAAACCAATTAAGGATGACAAAGATAGTCTTCTTAATGA 3791
3723 QY GAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTATGTTAGTCAATTTTCA 3782
3792 Db GAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTATGTTAGTCAATTTTCA 3851
3783 QY CTGCCCCATCAAAATTCAGATTCCTCTCTCAGCAGCAATTTTATGAGCTGCTGTA 3842
3852 Db CTGCCCCATCAAAATTCAGATTCCTCTCTCAGCAGCAATTTTATGAGCTGCTGTA 3911
3843 QY GTTCTTGTGACTTTCTTCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3902
3912 Db GTTCTTGTGACTTTCTTCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3971
3903 QY GAAGATCAAAACAAAGTTGTTGGGCTCTCAGACGTAGAGAGCAACTTCTTCAGAGATGCA 3962
3972 Db GAAGATCAAGCAGAGCTGTTGGGCACTCCAGAGAGAGAGCAACTTCTTCGGAGATGCA 4031
3963 QY ACAGATGGCAGCGCTCCCTTTGCTCTGTAAATGTGCTTGGAAACAGATGAGAGCC 4022
4032 Db ACAGATGGCAGCGCTCCCTTTGCTCTGTAAATGTGCTTGGAAACAGATGAGAGCC 4091
4023 QY TCCTGATCTTATTTGGGGAGTATAAGACTGTTCCCAACCACTTGCATGAGAGCCGTG 4082
4092 Db TCCTGATCTTATTTGGGGAGTATAAGACTGTTCCCAACCACTTGCATGAGAGCCGTG 4151
4083 QY TTTTGGCAACAAAGCCCT 4142
4152 Db TTTTGGTAAACAAAGCT 4211
4143 QY CATCCCT 4202
4212 Db AATCCCT 4271
4203 QY GCAGATCCGATAGTGTACAAAGGAGAGTCAAGGCGCTGAGAAACCGGAGAGCAGCAGCC 4262
4272 Db GCAGATCCGATAGTGTACAAAGGAGAGTCAAGGCGCTGAGAAACCGGAGAGCAGCAGCC 4331
4263 QY CCCTGCAAGCTGGGACCTGCACTCA 4290
4332 Db CCCTGCAAGCTGGGACCTGCACTTGA 4359

```

RESULT 8
AF119821
LOCUS          4313 bp mRNA linear ROD 16-APR-1999
DEFINITION    Mus musculus attractin (Mgca) mRNA, complete cds.
ACCESSION     AF119821
VERSION       AF119821.1 GI:4585306
KEYWORDS
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 4313)
AUTHORS       Gunn, T.M., Miller, K.A., He, L., Hyman, R.W., Davis, R.W., Azarani, A.,
Schlossman, S.F., Duke-Cohan, J.S. and Barsh, G.S.
TITLE         The mouse mahogany locus encodes a transmembrane form of human
attractin
JOURNAL       Nature 398 (6723), 152-156 (1999)
MEDLINE       99184160
PUBMED        10086356
REFERENCE     2 (bases 1 to 4313)
AUTHORS       Gunn, T.M., Miller, K.A., He, L., Hyman, R.W., Davis, R.W., Azarani, A.,
Schlossman, S.F., Duke-Cohan, J.S. and Barsh, G.S.
TITLE         Direct Submission
JOURNAL       Submitted (12-JAN-1999) HHMI, Stanford, Beckman Center B271A,
Stanford, CA 94305-5323, USA
FEATURES
Source
1..4313
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosomes="2"
1..4313
/feature="Mgca"
/feature="mahogany candidate"
27..4313
/feature="Mgca"
/feature="mahogany"
/codon_start=1
/product="attractin"
/protein_id="AAD25372.1"
/db_xref="GI:4585307"
/translation="MVAVAASFEARLGGSTTATAPAGRGKQHRPCTATGAWRG
PRARLCLPRVLRALPPPLPLPSLLLPREAAVAASAAVAASAAEKECD
PCVNGRCNPGQCVCTGWGQCCQCHGGRFLTGTSSGFVTDGPNKYKTKTCL
IEGYPNVLRLRFNFHATECSWDHLYVYDGSIIYAPLIAFSLIVPERDGNTEP
TVGYSALLHPFSDAAYNLGTGNITYNFDMCPNCSARGCKSGNSGSAVECEGSE
KGESCDI PHCTDNCGFPHRGICNASDTRGSCFPHWGPQCSI PVPANQSFWTREYS
DLKLPRASHKAVNGNIMVVGGMFNSHDSMLVLDLTSREWLPLNHSVNSVRY
GHSALHKDKIYMGCKIDSTGNVNEURVFIHNSWLLTPKAKDQIYAVGHSR
VTLASRVMLVIFGCEPLYGISVVOYDLEKNTWILHTQGLVQGVGHSSADD
RTKALYVHGGYKAFSANKYRLADDLRYDDVTQMTILKDSRFFRYLHTAVISG
VFGNTHDNTSGHAKCFSSDFWADYIADRWVLPRLPHLDVNFHGFSAVLNLT
MYVFGFNSLLSLDLVFTSQDCAHREAAVAGPGRICLWDTSRHSCTEWELATE
EQAEKLSKESKRTLDHRCDOHTDCYSTANTNDCHWNCNDCVPHNSCTEQI SI
AYESECPRDNPMYCNKTSRSCALDONCOWEPNQCIALPENI CNGWHLVGN
LKRTAKENYDNALSKRNHNLASLTQKVLVLKQLRMOSSQSMKLTUTPW
GLRKNVSWCWEDESPNLSLLOWPSEPDAGFCGLLSEPTRGKLAATCINFLNG
SVCEPANHSAKQCTPCALATAGCECTSSSECWMCNMQCVDNAYVASPFPGQ
MEWYTMSSCPENCSGYCTCHCLPQCGCCTDPSTNGKQCI EGSYKGPVMPQSA
SAGNVPOLPLNLSMCLDSRYNWSFTHCPACQCNHSCKINQSVCKEDLTGKH
ETCISGFYDPTNGKQCPCKNGHASICNTNTGKCTCTTKGVKGDSCOLCKVNRVQ
GNPLKGTCTYILLIDYQFTSLSQDDRYITAINFVATPDEQRDPMFINASKFNL
NLTWATSPAGTGTGEEVPVSKNIKESKDFSRNHPNITFFVIVSNFTWP
IKIQIAFSQHNFMDLVQFFVTFSCFLSLLLVAAVVMKIQSCWASRRQDLREHQ
QNASRPFASVNALETDEPDLLIGSITKPKPIALEPCFGKNAALVSVFRLPGL
GGIPPPGQGLAVASALVDISQOMPVIYVKEKSGAVNRNKOOPPAQPGTCI"
ORIGIN
Query Match      82.2%; Score 3527.2; DB 10; Length 4313;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 3821; Conservative 0; Mismatches 458; Indels 9; Gaps 1;
3 GGTGGCCGACGGCGCAACTGAGCAAGGCTGAGGAGGAGGAGCGGCGGACGGCAGC 62

```


Qy	3303	CAATGAGGGAATGTGACCATGCAAGTGCATGCGCAGCGCTCTGTGCAACACCAA	3362
Db	3326	TAATGAGGGAATGTGACCATGCAAGTGCATGCGCAGCGCTCTGTGCAACACCAA	3385
Qy	3363	CACGGGCAAGTGTCTTGCACCAACCAAGGCGCTCAAGGGGACGAGTGCCAGCTATGTA	3422
Db	3386	CACGGGCAAGTGTCTTGCACCAACCAAGGCGCTCAAGGGGACGAGTGCCAGCTATGTA	3445
Qy	3423	GGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAATGTTATATCTCTTCTAT	3482
Db	3446	GGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAATGTTATATCTCTTCTAT	3505
Qy	3483	TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGCTTATACAGCTATCAA	3542
Db	3506	TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGCTTATACAGCTATCAA	3565
Qy	3543	TTTTGTGGCTACTCTGACGAAACAAACAGGATTTGGACATGTTCAATGCTCCAA	3602
Db	3566	TTTTGTGGCTACTCTGACGAAACAAACAGGATTTGGACATGTTCAATGCTCCAA	3625
Qy	3603	GAATTTCAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAACCGAGCTGGAGA	3662
Db	3626	GAATTTCAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAACCGAGCTGGAGA	3685
Qy	3663	AGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA	3722
Db	3686	AGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA	3745
Qy	3723	GAAGTTGATTTTCGAAACCAACCAATATCACTTCTTGTGTTATGTCAGTAATTTAC	3782
Db	3746	GAAGTTGATTTTCGAAACCAACCAATATCACTTCTTGTGTTATGTCAGTAATTTAC	3805
Qy	3783	CTGGCCCATCAAAATTCAGATTCGCTTCTCTCAGACACAGCAATTTATGGACTCGTACA	3842
Db	3806	CTGGCCCATCAAAATTCAGATTCGCTTCTCTCAGACACAGCAATTTATGGACTCGTACA	3865
Qy	3843	GTTCTTCGTGACTTTCTCAGTGTGTTCTCTCTCTTCTGCTGCTGCTGCTGCTGCTG	3902
Db	3866	GTTCTTCGTGACTTTCTCAGTGTGTTCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTG	3925
Qy	3903	GAAGATCAAAACAAAGTTGTTGGGCTTCCAGACGATAGAGACAACTTCTTCGAGACATCA	3962
Db	3926	GAAGATCAAAACAAAGTTGTTGGGCTTCCAGACGATAGAGACAACTTCTTCGAGACATCA	3985
Qy	3963	ACAGATGGCCAGCGCT	4022
Db	3986	ACAGATGGCCAGCGCT	4045
Qy	4023	TCTGATCTTATTTGGGGGAGTATAAGACTGTTTCCCAACCCATTTGACCTGGAGCGTG	4082
Db	4046	TCTGATCTTATTTGGGGGAGTATAAGACTGTTTCCCAACCCATTTGACCTGGAGCGTG	4105
Qy	4083	TTTTGGCAACAAAGCGCTGCT	4142
Db	4106	TTTTGGCAACAAAGCGCTGCT	4165
Qy	4143	CATCCCTCTCTCTGCGCAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4202
Db	4166	CATCCCTCTCTCTGCGCAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4225
Qy	4203	GCAGATGCTCGATGTTTCAAGGAGAGTCAAGGAGCGCTGAGAAACCGGAGAGCAGCC	4262
Db	4226	GCAGATGCTCGATGTTTCAAGGAGAGTCAAGGAGCGCTGAGAAACCGGAGAGCAGCC	4285
Qy	4263	CCCTGACAGCTGGGACCTGCAATCTGA	4290
Db	4286	GCCTGACAGCTGGGACCTGCAATCTGA	4313

RESULT 9

AF034957

LOCUS

DEFINITION

Homo sapiens secreted T cell activation protein Attractin

ACCESSION	AF034957	AF034957.1	GI:3676346	(attractin) mRNA, complete cds.
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				Homo sapiens (human)
REFERENCE				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS				1 (bases 1 to 3597) Duke-Cohan, J.S., Gu, J., McLaughlin, D.F., Xu, Y., Freeman, G.J. and Schlossman, S.F.
TITLE				Attractin (DpPT-L), a member of the CVB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions
JOURNAL				Proc Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
MEDLINE				98403658
PUBMED				9736737
REFERENCE				2 (bases 1 to 3597) Duke-Cohan, J.S., Gu, J., Ao, Z., McLaughlin, D.F., Freeman, G.J. and Schlossman, S.F.
AUTHORS				Direct Submission
TITLE				Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
JOURNAL				
FEATURES				
source				Location/Qualifiers
gene				1..3597
CDS				/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" 1..3597 /gene="attractin" 1..3597 /note="attractin" /note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; DpPT-L" /codon_start=1 /product="Attractin" /protein_id="AAC61902.1" /db_xref="GI:3676347" /translation="MVAAAATEARLRRTTAATAALAGSGGPHCVNGRCNPGTQCQ VCPAGVQCOCHGGRFLLTGGTGGVDTGPGNYKTKTWTWLGPNRIMRLRNFN FATECSDHLYVDGDSIYAPLVAAPGLIIPERDNETVPVATSGVALLHFFSDA AYNLGNTIYFDMCPNNGSRGCKISNSDTECESENWKEACDIPHTDNCG FPHRGICNSDVRGCGSPDQWQPGCVFPANQSFWRREYSNLKLPASHKAVNG NIMVGGVTFNEDSHYMLAYDLASREWLPLNRSVNNVRYGHSALYKDKIYWG KTDSTGNTNELRVFHIHESWLLTPKAEQYAVVGHSAHIVLKNRVVYLVIFG HCLYGIYINVOBYDLDKNTWSLLHTQALVOGGYCHSSVYDHRTRALVHGGYKAFS ANKRYLADLLRYVDVTOMWTILKDSRPFYRLHTAVIVSGTMLVFGGTHNDSHSG AKCFSDFWAYDIACDRWVLPRDSTWMTDLAIPAVLHNTMTVFGGFFNSLLSDI LVFTSQCDARHSEACLAAGFIRCVNWTGSSQISWALADDEBEELKSCFSEKRT LQHDRCQDHTCYCTANTNDHCNDRHSCSEGQISIFRYENCKPNMPYIC NKRTSCRCALPQNCQWEPNQCIALPENICIGHWLVGNSCLIKITAKENYDNAL FPRNHALLSLTQKVEFLKQLRIMQSSQMSKLTLPWVGLRKINYSVWEDM SPFTNSLLOWMPSEPSDAGFCILSEPSTRGLKAATCINPLNGSVCEPANHSAQCR TPCALATACDCTGSSCEWCMCNKQCVSNAYVASPFQOCMEWYTWTCPPNCS GYCTCSHCLBQPCGCTDPSNTGKCKIEGSKGPKVMPQAPTFGNFVQPLLNSM CLEDYSNWFHCPACQCNHSHKLNQSIKCKENLTGKHCETCISGFDYDPPNGG KQPCPCNCHASLNTNKTCTGCTGKGVGSDQCEVENRTOGNPLRGTCTYLLID YQFTFSEQDEDDRYTAINFVATPDEQNRDLDMFINASKFNLTITWASFSAGTQAG EEMPVVSKTNIKEYKDSFSENEKDFRNHENITFFVYVSNTFTWPIKIQVTEQ"

ORIGIN

Query Match	80.2%;	Score	3469.4;	DB	9;	Length	3597;
Best Local Similarity	99.2%;	Pred. No.	0;				
Matches	3507;	Conservative	0;	Mismatches	26;	Indels	2;
Gaps							
Qy	270	GTGGGGCTCAGCGCAGCGAGCGCCAGGAATGTGACCGGCGCTGTGTCAACGGGGTGC	329				
Db	48	GGCGGCGACGGCAGCGCTCGCGGAGGAGCGCGGCGCGACTGTGTCAACGGGGTGC	107				
Qy	330	CTCACAACCTGGACCGGCAGTGGTCTGCCCGCGCTGGTGGGGAGCATGCCA	389				

108 Db CTGCAACCCCTGGCACCGGCCAGTCGCTGCTGCCCGCGGCTGGGTGGCGGAGCAATGCCA 167
390 Qy GCATCGCGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTCTGACAGATGACACTGG 449
168 Db GCATCGCGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTCTGACAGATGACACTGG 227
450 Qy AAAATTATAAATACAAAAGAAAGTGAACGTGGCTCATTTGAGGACAGCAATAGAAATAAT 509
228 Db AAAATTATAAATACAAAAGAAAGTGAACGTGGCTCATTTGAGGACAGCAATAGAAATAAT 287
510 Qy GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGA 569
288 Db GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGA 347
570 Qy TGGGACATCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCTCATTTGCTGAGAG 629
348 Db TGGGACATCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCTCATTTGCTGAGAG 407
630 Qy AGATGGCAATGAGACTGCTGAGTTGTTGCCACATCAGGTTATGCTTGTGCTCAATTT 689
408 Db AGATGGCAATGAGACTGCTGAGTTGTTGCCACATCAGGTTATGCTTGTGCTCAATTT 467
690 Qy TTTTATGTAATGCTGCTTAAATTTGACTGGATTTTAAATTAATTAATTAATTAATTAATG 749
468 Db TTTTATGTAATGCTGCTTAAATTTGACTGGATTTTAAATTAATTAATTAATTAATTAATG 527
750 Qy TCCAAATTAATGCTGACCGCGAGAGAGTGAAGTCAAGTAAATAGCAGCAAACTGTGA 809
528 Db TCCAAATTAATGCTGACCGCGAGAGAGTGAAGTCAAGTAAATAGCAGCAAACTGTGA 587
810 Qy ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869
588 Db ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
870 Qy CTGTGCTTTTCTCAATCGAGGCAATGTGCAATTCAGAGTGAATGTCAGAGATGCTGCTG 929
648 Db CTGTGCTTTTCTCAATCGAGGCAATGTGCAATTCAGAGTGAATGTCAGAGATGCTGCTG 707
930 Qy CTCAGACTGGAGGTCCTGGATGTTTCACTGCTGACAGTCAACCACTCAATTTGGAC 989
708 Db CTCAGACTGGAGGTCCTGGATGTTTCACTGCTGACAGTCAACCACTCAATTTGGAC 767
990 Qy TCGAGAGGAATATTTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGCTCAATGG 1049
768 Db TCGAGAGGAATATTTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGCTCAATGG 827
1050 Qy AAACATTAATGCTGTTGAGGATATATGTTCAACCACTCAGATTAATACATGTTCT 1109
828 Db AAACATTAATGCTGTTGAGGATATATGTTCAACCACTCAGATTAATACATGTTCT 887
1110 Qy AGGATATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCTTCTGTGAACAATGTGCT 1169
888 Db AGGATATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCTTCTGTGAACAATGTGCT 947
1170 Qy TGTTAGATATGCTATCTTTGGCATATTAAGGATTAATTAATTAATTAATTAATTAATG 1229
948 Db TGTTAGATATGCTATCTTTGGCATATTAAGGATTAATTAATTAATTAATTAATTAATG 1007
1230 Qy AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTAATAGATC 1289
1008 Db AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTAATAGATC 1067
1290 Qy ATGGGTGTTGTTGACCCCTAAGGCAAGGAGAGTATGAGTGGTTGGCATCTGTGACA 1349
1068 Db ATGGGTGTTGTTGACCCCTAAGGCAAGGAGAGTATGAGTGGTTGGCATCTGTGACA 1127
1350 Qy CATTTGTTACTGAAGATGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409
1128 Db CATTTGTTACTGAAGATGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187
1410 Qy CTATGGATATTAAGCAATGTGAGGAATATGATTTGGATAAGCAACACATGAGTATAT 1469
1188 Db CTATGGATATTAAGCAATGTGAGGAATATGATTTGGATAAGCAACACATGAGTATAT 1247

1470 Qy ACACACCCAGGGTGGCTTGTGCAAGGGGTTTACGGCCATAGCAGTGTGTTACGACCATAG 1529
1248 Db ACACACCCAGGGTGGCTTGTGCAAGGGGTTTACGGCCATAGCAGTGTGTTTACGACCATAG 1307
1530 Qy GACACGGGCGCTTATAGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1589
1308 Db GACACGGGCGCTTATAGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1367
1590 Qy TGAGATGATCTCTACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
1368 Db TGAGATGATCTCTACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1427
1650 Qy CCGATTTTCCGTTACTTGCACACACTGCTGATGATGATGATGATGATGATGATGATGATGATG 1709
1428 Db CCGATTTTCCGTTACTTGCACACACTGCTGATGATGATGATGATGATGATGATGATGATGATG 1487
1710 Qy GGAACACACACACACACATCTATAGGCAATGCGGCGCAATGCTTCTCTTCAGATTT 1769
1488 Db GGAACACACACACACACATCTATAGGCAATGCGGCGCAATGCTTCTCTTCAGATTT 1547
1770 Qy CATGGCTATGACATTTGCGCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1829
1548 Db CATGGCTATGACATTTGCGCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606
1830 Qy TGATGTCAAACAGATTTGGCCATT - CAGCAGTCTTTACACAAACAGCAACATGATGTTGCTG 1888
1607 Db TGATGTCAAACAGATTTGGCCATTCCAGCAGTCTTTACACAAACAGCAACATGATGTTGCTG 1666
1889 Qy GTGGTTTCAATAGTCTCTCTCAGCGACATCTGCTGATTTCACTCGAACAGTGTGATG 1948
1667 Db GTGGTTTCAATAGTCTCTCTCAGCGACATCTGCTGATTTCACTCGAACAGTGTGATG 1726
1949 Qy CGCATCGAGTGAAGCGCTTGTGTAGCAGCAGACCTGCTGATTTGCTGCTGCTGCTGCTGCTG 2008
1727 Db CGCATCGAGTGAAGCGCTTGTGTAGCAGCAGACCTGCTGATTTGCTGCTGCTGCTGCTGCTG 1786
2009 Qy CAGGTCGTCTCAGTGTATCTCGTGGCGCTGCGCACTGATGATGATGATGATGATGATGATG 2068
1787 Db CAGGTCGTCTCAGTGTATCTCGTGGCGCTGCGCACTGATGATGATGATGATGATGATGATG 1846
2069 Qy AATCAGATGCTTTTCCAAAGAACTCTTGCACATGACAGATGCTGACACAGACACACAGAT 2128
1847 Db AATCAGATGCTTTTCCAAAGAACTCTTGCACATGACAGATGCTGACACAGACACACAGAT 1906
2129 Qy GTTACAGCTGTACAGCCCAACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2188
1907 Db GTTACAGCTGTACAGCCCAACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1966
2189 Qy GGAACACAGCTGCTGAGAGGCGAGATCTCCATTTTGTAGTATGAGAAATGCTGCTGCTGCTG 2248
1967 Db GGAACACAGCTGCTGAGAGGCGAGATCTCCATTTTGTAGTATGAGAAATGCTGCTGCTGCTG 2026
2249 Qy ATAAACCCCATGTACTACTGTAAACAAAGAAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2308
2027 Db ATAAACCCCATGTACTACTGTAAACAAAGAAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2086
2309 Qy ACTGCGAGTGGAGCGCCCGGAAATCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2368
2087 Db ACTGCGAGTGGAGCGCCCGGAAATCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2146
2369 Qy TTGCTG 2428
2147 Db TTGCTG 2206
2429 Qy ACAATGCTTAAATTTCTGTAGGAAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2488
2207 Db ACAATGCTTAAATTTCTGTAGGAAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2266
2489 Qy AGAAGTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2548
2267 Db AGAAGTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2326

QY	2549	AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCTTACTGTGTCGTGGG	2608
DB	2327	AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCTTACTGTGTCGTGGG	2386
QY	2609	AAGATATGTCCCATTTTACAATAGTTTACTACAGTGGATGCCCTCTGAGCCCGAGTCATG	2668
DB	2387	AGATATGTCCCATTTTACAATAGTTTACTACAGTGGATGCCCTCTGAGCCCGAGTCATG	2446
QY	2669	CTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCTGGGGACTCGAAGCTTCGAACCTGCA	2728
DB	2447	CTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCTGGGGACTCGAAGCTTCGAACCTGCA	2506
QY	2729	TCAACCCCACTCAATGGTAGTCTCTGTGAAAGGCTCGAAACCAACAGTGTCTAAGCAGTGCC	2788
DB	2507	TCAACCCCACTCAATGGTAGTCTCTGTGAAAGGCTCGAAACCAACAGTGTCTAAGCAGTGCC	2566
QY	2789	GGACACCAATGTGCCCTTGAGGACAGCATGTGGAGATTGCACACAGCGGCAGCTCTGAGTGCA	2848
DB	2567	GGACACCAATGTGCCCTTGAGGACAGCATGTGGAGATTGCACACAGCGGCAGCTCTGAGTGCA	2626
QY	2849	TGTGGTGCAGCAACATGAGCAGTGTGTGACTTCAATGTGCTATGTGGCCTCTCTTCCCTT	2908
DB	2627	TGTGGTGCAGCAACATGAGCAGTGTGTGACTTCAATGTGCTATGTGGCCTCTCTTCCCTT	2686
QY	2909	TTGGCCAGTGTATGGAAATGATACGATGAGCACTGCCCCCTGAAATTTGTTCAAGGCT	2968
DB	2687	TTGGCCAGTGTATGGAAATGATACGATGAGCACTGCCCCCTGAAATTTGTTCAAGGCT	2746
QY	2969	ACTGTACTGTAGTCAATGTGTTGAGCAACAGAGCTGTGCTGGTGTACTGTATCCAGCA	3028
DB	2747	ACTGTACTGTAGTCAATGTGTTGAGCAACAGAGCTGTGCTGGTGTACTGTATCCAGCA	2806
QY	3029	ATACTGGCAAGGAAATGCAATAGAGGTTCCCTATAAGGACCAAGTCAAGATGCTTCCG	3088
DB	2807	ATACTGGCAAGGAAATGCAATAGAGGTTCCCTATAAGGACCAAGTCAAGATGCTTCCG	2866
QY	3089	AAGCCCTTACAGGAAATTTCTATCCACAGCCCCCTGTCTCAATTCAGCATGTGTCTAGAG	3148
DB	2867	AAGCCCTTACAGGAAATTTCTATCCACAGCCCCCTGTCTCAATTCAGCATGTGTCTAGAG	2926
QY	3149	ACACGAGATCAACTGTGCTCTTCAATCACTGTCCAGCTTCCAAATGCAAGCGGCACAGTA	3208
DB	2927	ACACGAGATCAACTGTGCTCTTCAATCACTGTCCAGCTTCCAAATGCAAGCGGCACAGTA	2986
QY	3209	AATGCATCAATCAGAGCATCTGTGAGAAATGTGAGAACTTGACACAGGCAAGCACTGG	3268
DB	2987	AATGCATCAATCAGAGCATCTGTGAGAAATGTGAGAACTTGACACAGGCAAGCACTGG	3046
QY	3269	AGACCTGCATATCTGGCTTCTAAGGTATCCCAACCAATGAGGGGAAATGTCAAGCCATGCA	3328
DB	3047	AGACCTGCATATCTGGCTTCTAAGGTATCCCAACCAATGAGGGGAAATGTCAAGCCATGCA	3106
QY	3329	AGTGCATGGGACGCGCTCTGTGTGAACACCAACACGGGCAAGTCTTCTGCACCAACA	3388
DB	3107	AGTGCATGGGACGCGCTCTGTGTGAACACCAACACGGGCAAGTCTTCTGCACCAACA	3166
QY	3389	AGGCGGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC	3448
DB	3167	AGGCGGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC	3226
QY	3449	CTCTCAGAGAAACATGTTATTATCTTCTTATTCAGTATCGATTCAGTTTACCTTTAGTCTAT	3508
DB	3227	CTCTCAGAGAAACATGTTATTATCTTCTTATTCAGTATCGATTCAGTTTACCTTTAGTCTAT	3286
QY	3509	CCGAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTTCAGCAACAAA	3568
DB	3287	CCGAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTTCAGCAACAAA	3346
QY	3569	ACAGGGAATTTGGACATGTTCAATGCGCTCCAGAAATTTCAACCTCAACATCACTGGG	3628
DB	3347	ACAGGGAATTTGGACATGTTCAATGCGCTCCAGAAATTTCAACCTCAACATCACTGGG	3406
QY	3629	CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGATGCTGTGTGTTTCAAAAACCA	3688

[illegible]

QY 411 ----- 410
Db 613 TGAGCGAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTGACTAAATTTTCCACC 672
QY 411 ----ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAATCAAAA 466
Db 673 GAGNACTAACTGGCTCTTCTGGATTTGTACAGATGGACCTGGGAATTAATAATAAGA 732
QY 467 CGAAGTGCAAGTGGCTCAATTGAAGGACAGCAAAATAGAAATATGAGACTTGTGTTCAATC 526
Db 733 CGAAGTGCAATGGCTCAATTGAAGGACAGCAAAATAGAAATATGAGACTTGTGTTCAATC 792
QY 527 ATTTTGTCTACAGAGTGTAGTGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586
Db 793 ATTTTGTCTACAGAGTGTAGTGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 852
QY 587 CAGCGTGTGTGCTATGATGAGTGGCTCATTTGCTTCCAGAGAGATGGCAATGAGACTG 646
Db 853 CAGCTGTGATGCTGCTTGTAGTGGCTCATTTGCTTCCAGAGAGATGGCAATGAGAGCG 912
QY 647 TCCTGAGGTTGTGGACATCAGGTTATGCCCTGTGCTGCATTTTATGATGGGACTCAATTTATG 706
Db 913 CTCTGAGGTCATGTCATTTAGGTTATGCACTGCTGCTGCTTTTTCAGTGAAGTGTGCTT 972
QY 707 ATAATTTGACTGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 766
Db 973 ATAATCTGACTGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1032
QY 767 GCCGAGGAGTGTAGATCAGTAAATAGCAGGCAACTGTTGAAATGTAATGTTCTGAAA 826
Db 1033 GCCGAGGAGTGTAGATCAGTAAATAGCAGGCAACTGTTGAAATGTAATGTTCTGAAA 1092
QY 827 ACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATC 886
Db 1093 ACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATC 1152
QY 887 GAGGCATCTGCAATTCAGTGTACAGAGTGTCTGCTTCTCAGACTGGCAGGTC 946
Db 1153 GAGGCATCTGTAATGACAGGATACAGAGGTTCTCTGCTTCTCAGTGGCAGGTC 1212
QY 947 CTGATGTTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1006
Db 1213 CTGATGTTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
QY 1007 ACTTAAAGCTCCCGAGGATCTCATAAAGCTGTGGTCAATGGAACCAATTTATGCTGTTG 1066
Db 1273 ATTTAAAGCTCCCGAGGATCTCATAAAGCTGTGGTCAATGGAACCAATTTATGCTGTTG 1332
QY 1067 TTGAGGATATATGTTCAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGCTT 1126
Db 1333 TTGGCGATATATGTTCAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGCTT 1392
QY 1127 CTAGGAGTGGCTCCACTAAACCTGTTGTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1186
Db 1393 CTAGGAGTGGCTCCACTAAACCTGTTGTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1452
QY 1187 CTTTGGCATTTACAGGATATAATTTACATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1246
Db 1453 CTTTGGCATTTACAGGATATAATTTACATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1512
QY 1247 ATGTGACCAATGATGAGGTTTTCATTTCAATTAATGAGTCATGGGTGTTGTTGACCC 1306
Db 1513 ACGTGACCAATGAGTGAAGTATTTTCATTTCAATTAATGAGTCATGGGTGTTGTTGACCC 1572
QY 1307 CTAAGGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCAACATTTGTTACACTGAAGA 1366
Db 1573 CGAAAGCTAAGGATCAGTATGAGTGGTGGGCACTCTGCAACATTTGTTACACTGGCAT 1632
QY 1367 ATGCCAGTGTGTCATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1426
Db 1633 CTGGCGTGTGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1692
QY 1427 ATGTGAGGAAATATGATTTGGATTAAGAACACATGAGTATATTACACACCCAGGTTGCC 1486

Db 1693 TTGTGAGGAATATGACTTTGGAAGAAACACATGAGTATATTACATATCTCAGGTTGCTC 1752
QY 1487 TTGTGCAAGGGGTTTACGCCCATAGCAGTGTATTCAGACCATAGGACAGGAGGCTTATACG 1546
Db 1753 TTGTGCAAGGGGTTTATGSCCAGTGTATTTATGATGACAGGACCAAGGCTCTGTACG 1812
QY 1547 TTCAATGGCTTACAGGCTTTTCAAGTGTGCAATTAAGTACCGGCTTGCAGATGATCTTACT 1606
Db 1813 TTCAATGGCTTACAGGCTTTTCAAGTGTGCAATTAAGTACCGGCTTGCAGATGATCTTACT 1872
QY 1607 GATATGATGTGATACCCAGATGTGACCATTTCTTAAAGACAGCCGATTTTTCGGTTACT 1666
Db 1873 GATACGATGTGATACTCAGATGTGCAATTTCTTAAAGACAGCCGATTTTTCGGTTACT 1932
QY 1667 TGACACAGCTGTGATGAGTGGAAACATGCTGGTGTGTTGGGGGAAACACACAAATG 1726
Db 1933 TGACACAGCTGTGATGAGTGGAAACATGCTGGTGTGTTGGGGGAAACACACAAATG 1992
QY 1727 ACACATCTATGAGCCATGGCGCAATGCTTCTTCTTCAATGATTTTTCAGATTTTTCAGATTT 1786
Db 1993 ACATTTCCATGAGCCAGGTTGCAATTTCTTCTCGACTTCTCGACTTATGACATG 2052
QY 1787 CTTGTGACCGCTGTGCTGCTTCCAGACCTGATCTTCCAGCATGATGTCAACAGATTTG 1846
Db 2053 CTTGTGACCGCTGTGCTGCTTCCAGACCTGATCTTCCAGCATGATGTCAACAGATTTG 2112
QY 1847 GCCATTTACAGCTTTTACACACAGCAGCATGATGTGTTGGTGTGTTTCAATGATCTTCC 1906
Db 2113 GCCATTTACAGCTTTTACACACAGCAGCATGATGTGTTGGTGTGTTTCAATGATCTTCC 2172
QY 1907 TCCTGAGCACTCTGCTGTTTACCTCGAAACAGTGTGCTGCGCATCGGAGTGAAGCG 1966
Db 2173 TCCTGAGCACTCTGCTGTTTACCTCGAAACAGTGTGCTGCGCATCGGAGTGAAGCG 2232
QY 1967 CTTGTTTACAGCAGGACCTGCTGTTTTCGGTGTGTTGGAACACAGGCTGCTCAGTGTG 2026
Db 2233 CTTGTTGTCAGCAGGACCTGCTGTTTTCGGTGTGTTGGAACACAGGCTGCTCAGTGTG 2292
QY 2026 TCTGCTGGGCGTGGCACTGATGAACAAAGAAAGAAAGTTAAATCAGAAATGTTTTTCTA 2086
Db 2293 CTTGCTGGGAGTTGGCACTGAAAGAAACAGCAGAAAGTTAAATCAGAAATGTTTTTCTA 2352
QY 2087 AAAGAACTCTTGACCATGACATGTGACAGCAGCAGATTTGTTACAGCTGTACAGCA 2146
Db 2353 AAAGAACTCTTGACCATGACATGTGACAGCAGCAGATTTGTTACAGCTGTACAGCA 2412
QY 2147 ACACCAATGACTGCCACTGGTGCATGACCAATGCTGTTGCCAGGAAACACAGCTGCTCAG 2206
Db 2413 ATACCAATGACTGCCACTGGTGCATGACCAATGCTGTTGCCAGGAAACACAGCTGCTCAG 2472
QY 2207 AAGGCCAGATCTCCATTTTGTGATGAGATTTGCCCAAGGATTAACCCCATGTTACTACT 2266
Db 2473 AAGGCCAGATCTCCATTTGCAAGTATGAGATTTGCCCAAGGATTAACCCCATGTTACTACT 2532
QY 2267 GTAAAGAGGACCACTGCTGAGGAGCTGTGCCCTTGGACCAAGCTGCCAGTGGGAGCCCC 2326
Db 2533 GCAATTAAGAAACCACTGCTGAGGAGCTGTGCCCTTGGACCAAGCTGCCAGTGGGAGCCCC 2592
QY 2327 GGAATCAGAGTGTGCTGCTGCCGAAATATCTGTTGGAATGCTGCTGCTGCTGCTGCTGCTGCT 2386
Db 2593 GGAATCAGAGTGTGCTGCTGCCGAAATATCTGTTGGAATGCTGCTGCTGCTGCTGCTGCTGCT 2652
QY 2387 GAACTCATGTTTGAATTTACTTCTGCAAGGAGATTTATGCAATGCTAAATTTGTTCT 2446
Db 2653 GAACTCATGTTTGAATTTACTTCTGCAAGGAGATTTATGCAATGCTAAATTTGTTCT 2712
QY 2447 GTAGGAACCAATGCTGCTTGTGCTTCTTACAAACCAAGAGAGGAGTGAATTTGCTC 2506
Db 2713 GTAGGAACCAATGCTGCTTGTGCTTCTTACAAACCAAGAGAGGAGTGAATTTGCTC 2772
QY 2507 TTAAGCAGCTGGCAATTAATGCAATGCTCAGAGCATGCTCAAGCTCACTTAACCCCAT 2566

Db 2773 TTAAGCAGCTTCGATTAAATGCAATCATCTCAAAAGTATGTCCAAGCTCACTCTGACTCCAT 2832
Qy 2567 GGGTCGGCCCTTCGGAAGATCAATGTCTCCCTACTGCTGGGAGATATGTCCCAATTTA 2626
Db 2833 GGGTGTGCTTCGGAAGATCAATGTCTCCCTACTGCTGGGAGATATGTCTCCCAATTTA 2892
Qy 2627 CAATAGTTTACTACAGTGATGCGCTCTGAGCCAGTGTGATGCTGTGGAATTT 2686
Db 2893 CAATAGTTTGTGCGATGATGCCATCTGAGCCAGTGTGATGCTGTGGAATTT 2952
Qy 2687 TATCAGAACCCAGTACTCGGGAGTGAAGCTGCAACCTGATCAACCCACTCAATGTTA 2746
Db 2953 TGTACAGCCCTAGTACTCGGGAGTAAAGCTGCAACCTGATCAACCCCTCTCAATGGCA 3012
Qy 2747 GTGTCTGTGAAGCCCTGCAAAACCAAGTGTCAAGCAGTGCAGGACCAACCATGTGCTTGA 2806
Db 3013 GCGTCTGTGAAGCCCTGCAAAACCAAGTGTCCAGCAGTGCAGGACCAACCATGTGCTTGC 3072
Qy 2807 GGCACAGATGTGAGATTTGCAACAGCGGAGCTCTGAGTGCAATGTGTCAGCAACATGA 2866
Db 3073 GGCACAGCTGTGCGGAGTGCATCTAGCAGCAGCTCGAGTGCATGTGTCAGTAACATGA 3132
Qy 2867 AGCAGTGTGAGTCCCAATGCTATGTGCGCTCTCTTCCCTTTGCGCAGTGTATGGAAT 2926
Db 3133 AGCAGTGTGAGTCCCAATGCTAGTGCCTCTCTTCCCTTTGCGCAGTGTATGGAAT 3192
Qy 2927 GGTATACGATGAGCCTCTCCCTCCCTGAAATTTGTTACGGTACTGTACTCTGTAGTCAAT 2986
Db 3193 GGTATACGATGAGCCTCTCCCTCCCTGAAATTTGTTCTGGTACTGTACTCTGCAGCCAT 3252
Qy 2987 GCTTGGAGCAACAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3046
Db 3253 GCTTGGAGCAGCCAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3312
Qy 3047 GCATAGAGGTTTCTTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAT 3106
Db 3313 GTATTGAGGCGAGCTATAAGAGGCTGTGAGATGCGCTCACAGGCTCTGCAGGAATG 3372
Qy 3107 TCTATCCACAGCCCTGCTCAATTCAGCATGTGCTAGAGAGCAGCAGATACCACTGGT 3166
Db 3373 TGTATCCACAGCCCTTCTGAATCTCAGCATGTGCTAGAGAGCAGCAGATACCACTGGT 3432
Qy 3167 CTTTCATTCACCTGTCCAGCTTCCCAATGCAAGCCAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAT 3226
Db 3433 CTTTCATTCACCTGTCCAGCTTCCCAATGCAAGCCAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATG 3492
Qy 3227 TCTGTGAGAGTGTGAGAACCTGACCAAGCAGGCTGCGAGACCTGCAATCTGCT 3286
Db 3493 TCTGTGAGAGTGTGAGAACCTGACCAAGGCGAGCTGCGAGACCTGCAATCTGCT 3552
Qy 3287 TCTACGGTATCCCAACCAAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGT 3346
Db 3553 TCTATGGTGAACCGGCTAATGAGGCAATGTGAGGCAATGTGAGGCAATGTGAGGCAATGTGAGGCAATGT 3612
Qy 3347 CTTGTGCAACCAACCAAGGAGTGTGAGGCAATGTGAGGCAATGTGAGGCAATGTGAGGCAATGT 3406
Db 3613 CACTGTGCAACCAACCAAGGAGTGTGAGGCAATGTGAGGCAATGTGAGGCAATGTGAGGCAATGT 3672
Qy 3407 AGTCCAGCTATGTGAGTGAATATGATACCAAGGAGTGTGAGGCAATGTGAGGCAATGTGAGGCAATGT 3466
Db 3673 AGTCCAGCTATGTGAGTGAATATGATACCAAGGAGTGTGAGGCAATGTGAGGCAATGTGAGGCAATGT 3732
Qy 3467 ATTATCTCTCTTATGACTATCAGTTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3526
Db 3733 ACTATACCTTCTCTATGACTATCAGTTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3792
Qy 3527 ATTACAGCTATCAATTTTGTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3586
Db 3793 ACTACAGCCATCAATTTTGTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3852
Qy 3587 TCATCAATGCTTCCAAAGATTTCAACTCAACATCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 3646
Db 3853 TCATCAATGCTTCCAAAGATTTCAACTCAACATCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 3912

Qy 3647 GAAACCAGCTGGAGAGAGATGCTGTGTTTCAAAAAACAACATTAAGAGTACAAAG 3706
Db 3913 GAAACCAGCTGGAGAGAGATGCTGTGTTTCAAAAAACAACATTAAGAGTACAAAG 3972
Qy 3707 ATAGTTCTCTAATGAGAAGTTTGATTTTGGACCAACCAACCAATATACATTTCTTTGTTT 3766
Db 3973 ATAGTTCTCTAATGAGAAGTTTGATTTTGGACCAACCAACCAATATACATTTCTTTGTTT 4032
Qy 3767 ATGTAGTAATTTCACTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAAT 3826
Db 4033 ATGTAGTAATTTCACTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAAT 4092
Qy 3827 TTATGGACCTGGTACAGTTCTTCTGACTTTCTCAGTTGTTTCTTCTTCTTCTTCTTCTTCTTCT 3886
Db 4093 TCATGGACCTGGTACAGTTCTTCTGACTTTCTTCTGACTTTCTTCTGACTTTCTTCTGACTTTCTTCTG 4152
Qy 3887 TGGCTGCTGTGTTTGAAGATCAAAAGTTTGTGGGCTCCAGACGTAGAGAGCAAC 3946
Db 4153 TGGCTGCTGTGTTTGAAGATCAAAAGTTTGTGGGCTCCAGACGTAGAGAGCAAC 4212
Qy 3947 TTCTTCAGAGATGCAACAGATGCGCAGCGCTTCTTCTGCTCTGTAATATGCTGCTTCTTCTGCTTCT 4006
Db 4213 TTCTTCAGAGATGCAACAGATGCGCAGCGCTTCTTCTGCTCTGTAATATGCTGCTTCTTCTGCTTCT 4272
Qy 4007 AAACAGATGAGGAGCTTCTGATCTTATTTGGGGGAGTATAAGACTGTTTCCCAACCCA 4066
Db 4273 AAACAGATGAGGAGCTTCTGATCTTATTTGGGGGAGTATAAGACTGTTTCCCAACCCA 4332
Qy 4067 TTGCTGAGGAGCTGTTTGGCAAAAGCCGCTCTCTCTCTGTTTGTGAGCTCC 4126
Db 4333 TTGCTGAGGAGCTGTTTGGCAAAAGCCGCTCTCTCTCTGTTTGTGAGCTCC 4392
Qy 4127 CTGAGGCTGCTGCTGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4186
Db 4393 CTGAGGCTGCTGCTGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4452
Qy 4187 TGTGACATTTCTCAGCAGATGCGGATAGTGTACAGGAGAGTTCAGGAGCGTGAGAA 4246
Db 4453 TGTGACATTTCTCAGCAGATGCGGATAGTGTACAGGAGAGTTCAGGAGCGTGAGAA 4512
Qy 4247 ACCGAGAGCAGCAGCCCTGCGACAGCTGGACCTGCATCTGA 4290
Db 4513 ACCGAGAGCAGCAGCCCTGCGACAGCTGGACCTGCATCTGA 4556

RESULT 11
AB038388
LOCUS
DEFINITION
Rattus norvegicus Atrn mRNA for attractin, complete cds.
ACCESSION
AB038388.1 GI:12275311
VERSION
AB038388.1
KEYWORDS
attractin.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (sites)
Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,
Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T.,
and Serikawa, T.
Attractin/mahogany/zitter plays a critical role in myelination of
the central nervous system
Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
21143347
MEDLINE
PUBMED
11209055
REFERENCE
2 (bases 1 to 4500)
Kuramoto, T., Serikawa, T. and Ushijima, T.
Direct Submission
Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan [E-mail:tkuramoto@gan2.ncc.go.jp]

Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)

FEATURES

source
1..4500
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="WTC"
/db_xref="taxon:10116"
1..4500
/gene="Atrn"
61..3888
/gene="Atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB21018.1"
/db_xref="GI:12275312"

gene

CDS

/translation="MVAAAAAEEALEARLGYTTATAPAGWKERQHRPCOATGAWRP
WPRAGLCPLRVLSRPPPLPLPLLSLLPLPREAEEAAVAASGSAEAAK
ECDPRVGRNCPGTGCVPTGWCQCGRPRLTGSSGFVDTGPGNYKTK
CTWLEGPKNILMRHNHPATECSWDHLVYDGDLSIYAPLIAFSGLIYPERDGNKT
APEVITSGYALLHFFSDAAYNLGTENITYNEDMPCNCSGRGECKSSNSSTVECEC
SENWKGSCDIPHCCTDNGCFPHRGICNASDTRGSCFPHQGGGCSIPVPANQSPWR
EYSDLKLPRASHAEVNGNIMVVGGMFNHSDYSWVLAYDLASREWSLNSVNSV
VVRVGHSLALHKDIXIMYGKIDSTGNVTNELRVFTHNESWLLTPKADQVAVGH
SAHVTLSGGRVLMVIFGHCPLYISVQVEYDEKRTWSIIQTQALVQGGYGHSS
YDHRTRKALYVHGQYKAFSANKYRDLADLRYHVDITQWTLIKDSRFRYLHTAVIS
GTMLVFGNTHNDTSMHGKACFSSDFMAYDIACDRWSVLPRPELHRDVRNRFHSAVL
HNSWYVPGFNLSLLDVLFTSEQCDHSEAAVAGPGIRCLWDTOSSRCTSWE
LATBEQAKLSECFKRLDHRDQDTCYCTANTNDCHNCNDHCVPNHESCTEG
QISIAKYNCPKPNMYCNKKTSCRSALDONCOWEPNCEGICALPENICGLWHLV
GNSCLKITAKENYDMSKRNHNAFLASITSKQKVEFVLKQRLMSSGOSTSKLTL
TPVGLRKNISYCWEDMSPTNLLQMPSEPSDAGFCGILSEPTRGILKATCIN
PLNGSVCRPANHSACQRTPCALRTACGECTSSSECSCNNKQCVDSNAYVAPFM
PGCWMEVMTSCPENCSGYCTCSHCLLEOPGCGWCTDPNNTGKGIKESYKGPVM
FSAHSTGNVYFQPLLNSMCLDSRYNWSFIHCPAQOCHGSKCINOSICEKCDLTT
KHCBETCISGFGDPTNGKOPKNCNGHSLCMTNNTGKFCPTTKYKGBECQCEVE
NRYQGNPLKGYCYLLIDYQFTFSLQEDDRYTAINFATVPDEQRDLMLINAK
NFNLTITWATSPFAGTQTGEEVPVVSKTNIKEYKDSFNEXEDFRNHNENITFFVYVSN
FTWPIKIQRVTS"

ORIGIN

Query Match 72.5%; Score 3111.6; DB 10; Length 4500;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 3369; Conservative 0; Mismatches 429; Indels 0; Gaps 0;
QY 3 GGTGCGCGCAGCGCGGCAACTGAGCAAGGCTGAGGAGGAGCAGCGCGGCGAGCGGAGC 62
DB 72 GCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 131
QY 63 GCTCGCGCGCAGGAGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 122
DB 132 GCGCGCGCGGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 191
QY 123 GGGGCTGGGGCGGGGCTGCGGCTCCCGGGCTGCTGCTCCACCGCTGCGGCGGCGGCT 182
DB 192 GTGGCGCGCGCGGGGCTGTGCTCCCGGGGCTGCTGCTCCCGGGGCTGCTGCTCCCGG 251
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
DB 252 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
QY 243 CGAGGCGCGCGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
DB 312 CGAGGCGCGCTGCGGCTGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 371
QY 303 TGACCGGCGCTGTGTCAAAGGCGGCTGCTGCAACCTGGCAGCGGCGGCTGCTGCTGCTG 362
DB 372 TGACCGGCGCTGTGTCAAAGGCGGCTGCTGCAACCTGGCAGCGGCGGCTGCTGCTGCT 431
QY 363 CGCGCGGCTGGTGGGCGAGCAATGCCAGACTCGCGGGGCGCGGCTTACAGACTACTGGATC 422
DB 432 CACGGGCTGGTGGGCGAGCAATGCCAGACTCGCGGGGCGCGGCTTACAGACTACTGGATC 491
QY 423 TTCTGGGTTTGTGACAGATGACCTTGGAAATTATAAATACAAACGAAAGTGACGCTGGCT 482

DB 492 TTCTGGATTGTAAACAGATGGACCTGGGAATTATAAATATAAAGCAATGACATGGCT 551
QY 483 CATTAAGAGCAGCGCAAAATAGAAATATGAGACTTCGTTTCATCATTTTCTACAGAGTG 542
DB 552 CATTAAGAGCAGCGCAAAATAGAAATATGAGACTTCGTTTCATCAACCAATTTTCTACGAATG 611
QY 543 TAGTTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATGACACCGCTAGTTGCTGC 602
DB 612 TAGCTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTTACGCACCTCTGATTGCTGC 671
QY 603 ATTAAGTGGCCTCATTTGCTCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGC 662
DB 672 CTTTACTGGCCTCATTTGCTCTGAAAGAGATGGCAATGAGACTGTCCTGAAAGTCACTGT 731
QY 663 CACATCAGGTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
DB 732 CACTTCAGGTTATGCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
QY 723 TAATATTACTTACAGTTTGTATATGTGTCCTCAATTAATCTGCTCAGGCGGAGGAGTGTTAA 782
DB 792 TAATATCACTTACATTTTGTGACATGTGCCGGAATAATGCTCAGGCGGAGGAGATGTAA 851
QY 783 GATCAGTAATAGCAGCGAACTGTTCAATGTGAAAGTCTGCTGAAAGCTGAAAGGTGAAGC 842
DB 852 GAGCAGTAACAGCAGCAGCAGCTGTTGAGTGTGAATGTTCTGAAAGCTGAAAGGGAGTTC 911
QY 843 ATGTGACATTCCTCCTCCTGACAGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 902
DB 912 CTGTGACATTCCTCCTCCTGACAGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971
QY 903 AAGTGAATGTCAGAGGATGCTCCTGCTCTCAGACTGCGAGGCTCCTGCTGCTGCTGCTGCTG 962
DB 972 AAGTGAACACAGAGGCTGCTCCTGCTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1031
QY 963 TGTACAGGCTAACAGGCTCAATTTTGGAGCTCGAGAGGAATATTTAACTTAAAGCTCCCGAG 1022
DB 1032 TGTGCCAGCTAACAGGCTCAATTTTGGAGCTCGAGAGGAATATTTCTGACTTAAAGCTCCCGAG 1091
QY 1023 AGCATCTCATAAAGCTGTGCTCAATGGAAACATTTATGCTGGGTTGTTGGAGGATATATGTT 1082
DB 1092 AGCTCTCATAAAGCTGAGCTCAATGGAAATATAATGCTGGGTTGTTGGCGGATATATGTT 1151
QY 1083 CAACCACTCAGATTATTAACATGCTTCTAGCGTATGACCTTCTTCTAGGAGGAGTGGCTTC 1142
DB 1152 CAACCACTCAGATTACAGCATGCTTTTGTAGCGTATGACCTGCTTCTAGGAGGATGGCTTC 1211
QY 1143 ACTAAACCGTCTGCTGAAACATGTTGTTGTAGATATGCTCATTTCTTGGCATTTACAA 1202
DB 1212 ACTAAACCATTTCTGTGACAGTGTGTTGTAGATATGCTCATTTCTTGGCATTTACAA 1271
QY 1203 GGATPAAATTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1262
DB 1272 GGATPAAATTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1331
QY 1263 GAGAGTTTTCACATTTCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1322
DB 1332 GAGAGTTTTCATTTTCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1391
QY 1323 GTATGCAAGTGGTGGGCACTCTGCAACATTTGTATCACTGAGAGGAGGAGGAGGAGGAGGAG 1382
DB 1392 GTATGCAAGTGGTGGGCACTCTGCAACATTTGTATCACTGAGAGGAGGAGGAGGAGGAGGAG 1451
QY 1383 CTG 1442
DB 1452 GTTGGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1511
QY 1443 TTGGAAGAAGAACACATGAGGATATATTAACACACCGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1502
DB 1512 CTTGGAAAGAAGAACACATGAGGATATATTAACAACTCAAGGCTGCTGCTGCTGCTGCTGCTG 1571
QY 1503 CGGCCATAGAGTGTGTTACGACCATAGGACACAGGCGGCTTACCTGCTGCTGCTGCTGCTGCT 1562

Db 1572 TGCCACAGCAGTGTATTATGATCACAGGACCAAGGCTCTGTATGTTCATGTGGCTACAA 1631
QY 1563 GGCTTTTCAGTGCCAAATTAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC 1622
Db 1632 GGCTTTTCAGGCCCAACAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC 1691
QY 1623 CAGATGTGGACCAATCTTAAGGACAGCGATTTTTCGGTTACTTGTGACACAGCTGTGAT 1682
Db 1692 TCAGATGTGGACCAATCTTAAGGACAGCGATTTTTCGGTTACTTGTGACACAGCTGTGAT 1751
QY 1683 AGTGAGTGGAAACCAATGCTGTGTTCGGGGGAAACACACACAATGACATCTATGAGCCA 1742
Db 1752 AGTGAGTGGAAACCAATGCTGTGTTCGGGGGAAACACACACAATGACATCTATGAGCCA 1811
QY 1743 TGGCGGCAAAATGCTTCTCTCAGATTTTCAGATTTTCAGGCTTATGACATTTGCGCTGTGACCGCTGTG 1802
Db 1812 CGGTGCAAAATGCTTCTCTCAGATTTTCAGATTTTCAGGCTTATGACATTTGCGCTGTGACCGCTGTG 1871
QY 1803 AGTGCTTCCAGACCTGATCTCCACATGATGTCACAGATTTGCGCAATTCAGAGCTTT 1862
Db 1872 AGTTCTCCCGACAGCTGAGCTCCATCAGATGTCAACCGATTTGCGCAATTCAGAGCTTT 1931
QY 1863 ACACAAACAGCACCATATGTGTTCGGTGGTTTCAATAGTCTCTCTCAGCGACATCCT 1922
Db 1932 GCACAAACAGCACCATATGTGTTCGGGCGCTTCAACAGCGCTCTCTCAGTGACGTCTT 1991
QY 1923 GGTATTCACCTCGGAACAGTGTGATCGCGATCGGAGTGAAGCGCTGTTTGAAGCAGG 1982
Db 1992 AGTCTTCACCTCGGAGCAGTGTGATGCAACCGCAGCGAAGCTGCTGTGTGGCAGCAGG 2051
QY 1983 ACTGTGATTCGGTGTGTGGAAACACAGGGTCGTCTCAGTGTATCTCGTGGGCTGCG 2042
Db 2052 ACCTGGCATCCGGTGTCTGTGGGACACACAGTCTCTCGATGTACTCTCGGAGTTGGC 2111
QY 2043 AACTGATGAACAGGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTTGACCA 2102
Db 2112 AACTGAAGAACAGCAGAAAGTTAAATCAGATGTTTTTCTAAAGAACCTTTGACCA 2171
QY 2103 TCACAGATGTGACAGCACACAGATTTGTACAGCTGTACAGCAACACCAATGATCTGCCA 2162
Db 2172 TGACAGATGTGACAGCACACAGATTTGTACAGCTGTGACAGCAATACCAATGATCTGCCA 2231
QY 2163 CTGGTGCAATGACCAATTTGTCGCCAGGAACACAGCTGTCTCAGAAGCCAGATCTCCAT 2222
Db 2232 CTGGTGCAATGATCACTGTGTCTGTGAAACACAGCTGTGACAGAGCCAGATCTCCAT 2291
QY 2223 TTTTAGGTATGAGATTTGCCCAAGGATAACCCATGTACTACTGTAAACAGAGACAG 2282
Db 2292 TGCCAGATGATGACAAATGGCCCAAGGATAACCCATGTACTACTGTAAACAGAGACAG 2351
QY 2283 CTGACAGAGCTGTGCCCTGGACCAAACTGCCAGTGGAGGCCCGGAATCAGGAGTGCAT 2342
Db 2352 CTGACAGAGCTGTGCCCTTAGATCAGAAATGGCAGTGGGAACCTCGAATCAGAGTGCAT 2411
QY 2343 TGCCCTTGCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCATGTTGAA 2402
Db 2412 CGCCCTGCCCCGAAACATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCATGTTGAA 2471
QY 2403 AATTACTACTGCCAGGAGAAATTTGACAAATGCTTAAATTTGTTCTGTAGGACCAATGC 2462
Db 2472 AATCACTACTGTAAAGAAATTTATGACAAATGCTTAAATTTGTTCTGTAGAAACCAATGC 2531
QY 2463 CCTTTTGGCTTCTTTTCAACCCAGAGAGAGTGAATTTGTCTTAAGCAGCTGCGAAT 2522
Db 2532 CTTTTTGGCTTCTTTTACCTCCAGAGAGAGTGAATTTGTCTTAAGCAGCTTGCAT 2591
QY 2523 AATGAGTCACTCAGAGCATGTCCAGCTCACTTAACCCATGGCTGCGCTTGGAA 2582
Db 2592 AATGCAATCATCTCAGAGCATGTCCAGCTCACTTCACTTCACTTCACTTCACTTCACT 2651
QY 2583 GATCAATGTGTCTTACTCTGGTGTGGAGATATGTCCTCCATTTACAAATAGTTTACTACA 2642
Db 2652 GATCAATGTATCTTACTCTGGTGTGGAGATATGTCCTCCATTTACAAATAGTTTACTACA 2711

QY 2643 GTGCATGCCGCTTGAGGCCAGTGTGATTTCTGTGGAAATTTTATCAGAAACCCAGTAC 2702
Db 2712 GTGCATGCCATCTGAGGCCAGTGTGATGCCGCTTCTGTGGATCTTGTGAGGCTTAGTAC 2771
QY 2703 TCGGGGACTGAAGGCTCAACCTGTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCC 2762
Db 2772 TCGGGGACTGAAGGCTCAACCTGTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCC 2831
QY 2763 TGCACCAACAGTGTGAAGGCTCGGACACCATGTGCTTGGAGCAGCATGTGGAGA 2822
Db 2832 TGCACCAACAGTGTGAAGGCTCGGACACCATGTGCTTGGAGCAGCATGTGGAGA 2891
QY 2823 TTGCACAGGCGGAGCTCTGAGTGCATGTGCTGCAGCAACATGAAGCAGTGTGTGACTC 2882
Db 2892 GTGCACAGCAGCAGCTCCGAATGATGTGTGCAGCAGCATGAAGCAGTGTGTGACTC 2951
QY 2883 CAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATATGATGAGCAG 2942
Db 2952 CAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATATGATGAGCAG 3011
QY 2943 CTGCCCCCTGAAATTTGTTTCAGGCTACTGTACTGTAGTCAATTTGTTGGAGCAACAGG 3002
Db 3012 CTGCCCCCTGAAATTTGTTTCAGGCTACTGTACTGTAGTCAATTTGTTGGAGCAACAGG 3071
QY 3003 CTGTGGCTGTGTACTGATCCCAAGCAATCTGCAAGGAAATATGATAGAGGTTCTTA 3062
Db 3072 CTGTGGCTGTGTACTGATCCCAAGCAATCTGCAAGGAAATATGATAGAGGTTCTTA 3131
QY 3063 TAAAGGACAGTCAAGATGCTTCCGAGCCCTCAGAGCAATTTCTATCCACGCCCCCT 3122
Db 3132 TAAAGGACAGTCAAGATGCTTCCGAGCCCTCAGAGCAATTTCTATCCACGCCCCCT 3191
QY 3123 GCTCAATTCAGCATGTGTCTAGAGCAGCAGATACAACTGTCTTTCTATCTACTGTCC 3182
Db 3192 TCTGAACTCCAGCATGTGTCTAGAGCAGCAGATACAACTGTCTTTTATCTACTGTCC 3251
QY 3183 AGCTTCCCAATGCAAGGCGCAGATGATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242
Db 3252 AGCTTCCCAATGCAAGGCGCAGATGATGATCAATCAGAGCATCTGTGAGAGTGTGA 3311
QY 3243 GAACTTGAACACAGGCAAGCAGCTCCGAGACCTGCAATCTGCTTCTACGGTATCCCAAC 3302
Db 3312 GGACTTGAACACAGGCAAGCAGCTCCGAGACCTGCAATCTGCTTCTATGGTGAACCCAG 3371
QY 3303 CAATGGAGGAAATGTGACCCATGCAAGTGCATGCGGACGCGTCTCTGTGCAACACCAA 3362
Db 3372 TAAATGGAGCAATGTGACCCATGCAAGTGCATGCGGACGCGTCTCTGTGCAACACCAA 3431
QY 3363 CACGGCAAGTGTCTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCTATGTGA 3422
Db 3432 CACTGGCAAGTGTCTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCTATGTGA 3491
QY 3423 GGTAGAAAATGATACCAAGGAAACCTCTCTCAGAGCAACATGTTATTTACTCTTTTAT 3482
Db 3492 GGTAGAAAATGATACCAAGGAAACCTCTCTCAGAGCAACATGTTATTTACTCTTTTAT 3551
QY 3483 TGACTATCAGTTTCACTTTTGTATCCCGAGGATGATGCTTATTTACAGATATCAA 3542
Db 3552 TGACTATCAGTTTCACTTTTGTATCCCGAGGATGATGCTTATTTACAGATATCAA 3611
QY 3543 TTTTGTGGCTACTCTCTGACGAAACAAAACAGGATTTGGACATGTTTCATCAATGCTCCAA 3602
Db 3612 TTTTGTGGCTACTCTCTGATGAAACAAAACAGGATTTGGACATGTTTCATCAATGCTCCAA 3671
QY 3603 GAATTTCACTCAGATCACTTGTATCCCGAGGATGATGCTTATTTACAGATATCAA 3662
Db 3672 AAATTTCACTCAGATCACTTGTATCCCGAGGATGATGCTTATTTACAGATATCAA 3731
QY 3663 AGAGATGCTGTGTGTTTCAAAAACCAACATTAAGGAGTATCAAGATAGTTTCTTAATGA 3722
Db 3732 AGAGATGCTGTGTGTTTCAAAAACCAACATTAAGGAGTATCAAGATAGTTTCTTAATGA 3791

```
Qy 3723 GAAGTTTGAATTTTCGCAACCAACCAATATACATCTTCTTGTGTTTATGTCAGTAATTTTCC 3782
Db 3792 GAAATTTTGAATTTTCGCAACCAATATACATCTTCTTGTGTTTATGTCAGTAATTTTCC 3851
Qy 3783 CTGGCCCATCAAAATTTCA 3800
Db 3852 TTGGCCCATCAAAATTTCA 3869

RESULT 12
AB093245 AB093245 5683 bp mRNA linear ROD 07-FEB-2003
LOCUS Mus musculus mRNA for "KIAA0548 protein.
DEFINITION AB093245
ACCESSION AB093245
VERSION AB093245.1 GI:26006172
KEYWORDS FLI_CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T.,
Ohara,O. and Koga,H.
Prediction of the coding sequences of mouse homologues of KIAA
gene: 1. The complete nucleotide sequences of 100 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
Unpublished
2. (bases 1 to 5683)
Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
Direct Submission
Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kanamari, Kisarazu, Chiba 292-0818, Japan
(E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
Location/Qualifiers
1..5683
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mb03041"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"
1..5683
/gene="KIAA0548"
<1..3181
/gene="KIAA0548"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=2
/evidence=not experimental
/product="KIAA0548 protein"
/protein_id="BAC1429.1"
/db_xref="GI:26006173"
/translation="AYDLTSREWLPLNHSVNSVYVYGHSLAHKDKIYMGKIDST
GNVNELRVFIHNSWLLTPKADQVAVGVSHIVTLASGRVYMLVIFGHCPLYG
YLSVQVEYLEKNTWSILHTQCALVQGGVGHSSVDDRTKALYHGVYKAFSANKYRL
ADLYRYDVTQMTWILKDSRFRVLHTAVIVSGTMLVFGGNTHDNTSMHGAKCFSS
DFNAYDIACDRWSLPRPELHDVNRFGHSVILVNSTMYVFGFNLSLLSDVLPTLE
QCDARSEAAVGAAGDRIECLMDTSSRCTSWELATEOEAKLSECKSEKSLDHDRC
DHTDCYCTANTDCHNCVFNHSCTEGOI SI AKYESCPKDNPMYCNKRTSC
RSCALDONQWEPNLCIALPENI CNGWHLVGNLSCLIT TAKENYDNKLSCRNTH
AFLATSTOKVYFVLKOLRMLQSSOSKSLTLPWVGLRKLINSVYQWEDMSPFTNS
LLQWMESEPSDAFCGILSEPSRGIKAATCINPLNGSVCPANESAKQCRTPCLAR
TAGCECTSSSECMWCSNNKQCVDSNAYVAFPPQOCMEWTWSSCPNCSGYCTCS
HCLQPGQGWCTDPSNTGKCI EGSYKGPVMPSPSASGNVYFQPLNNSNCLSDSR
YVNSWFTHCPCQCNKSCINOSVCEKEDLTFTGRHCETCISGFGDPNGKQCPFK
CNGHASLCLNTGCKFCFTTKGVKGDCEQLCEVENRYQGNPLKGT CYTLLLLIDYQFTFS
```

```
LSQEDRYTAINFVATPEBQNRD.LDMFINASKNPLNITWATSPAGTQGEVPVV
SKTTIKYKQDSFNEKFDERNHPNITFFVYNSFTWPIKIQIAFSQHNFDLVQFFV
TFPSCFLSLLVAVVWKIKOSCWASRRERQILREMQQWASRPFSVVALETDEFP
DLIGSIIKTVPKPIALEPCFGNKAALSVFVRLPRGLGIPPGQSGLAVALVDIS
QQMPIVYKEKSGAVNRKQQPPAQPGTCI"
ORIGIN
Query Match 62.8%; Score 2693; DB 10; Length 5683;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 2876; Conservative 0; Mismatches 305; Indels 0; Gaps 0;
Qy 1110 AGCGTATGACCTTGGTCTTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAAATGTGTT 1169
Db 1 AGCGTATGACCTTGGTCTTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAAATGTGTTGTT 60
Qy 1170 TGTTCATATGTTGTCATTTCTTGGCAATATATCAAGGATAAAATTTACATGTATGGAGAAA 1229
Db 61 TGTTCATATGTTGTCATTTCTTGGCAATATATCAAGGATAAAATTTACATGTATGGAGAAA 120
Qy 1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTC 1289
Db 121 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCATTAATGAGTC 180
Qy 1290 ATGCGTGTGTTGTTGACCCCTAAGGCAAGGAGAGTATGAGTGGTGGGCACTCTCCACA 1349
Db 181 AUGGTTATTTTAACTCCGAAAGCTAAGGATCAGTATGAGTGGTGGGCACTCTCCACA 240
Qy 1350 CATTGTTACACTGAAGAAATGGCCGAGTGGTCTATGCTGGTCACTTTGGTCACTGCCCTCT 1409
Db 241 CATTGTTACACTGGCACTCTGGCCGTGTGGTCTATGTTGGTCACTTTGGTCACTGCCCACT 300
Qy 1410 CTATGATATATTAAGCAATGTGCGAGGAATATGATTTGGATAAGAACACACATGAGTATAT 1469
Db 301 CTATGATATATTAAGCGTGTGCGAGGAATATGATTTGGATAAGAACACACATGAGTATAT 360
Qy 1470 ACACACCCAGGGTGCCTTGTGCAAGGGGTTACGGCCCATAGCAGTGTTTACGACCATAG 1529
Db 361 ACATACCTCAGGGTGTCTTGTGCAAGGGGTTATGCGCCACAGTAGTGTGTTATGATGACAG 420
Qy 1530 GACGAGGGCCCTATACGTTTCATGCTGGTCAAGGGTTCAGTGGCAATAGTACCGGCT 1589
Db 421 GACCAAGGCTCTGTACGTTTCATGCTGGTCAAGGGTTCAGGCGCCAAACAAATACCGCT 480
Qy 1590 TGCAAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGACCATCTTTAAGGACAG 1649
Db 481 TGCAAGATGATCTCTACAGATACGATGTGGATCTCAGATGTGACCATCTTTAAGGACAG 540
Qy 1650 CCGATTTTCCGTTTACCTTGACACACAGCTGTGATAGTGGTGGAAACCATGCTGTGTTGG 1709
Db 541 CCGATTTTCCGTTTACCTTGATACAGCTGTGATAGTGGTGGAAACCATGCTGTGTTGG 600
Qy 1710 GGGAACACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTTCTCTCAGATTT 1769
Db 601 AGGGAACACACAAATGACATCTTCAAGCCACCGTGGCCAAATGCTTCTCTCAGATTT 660
Qy 1770 CATGGCCTATGACATTTGCCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA 1829
Db 661 CATGGCTTATGACATTTGCTTGTGACCGATGGTGTGCTTCCAGACCTGAGTCTCCATCA 720
Qy 1830 TGATGTCAACAGATTTGGCCATTTCAGCAGTCTTACACACAGACCATGATGTGTTGG 1889
Db 721 TGATGTCAACAGATTTGGCCATTTCAGCAGTCTTGTACAACAGACCATGATGTGTTGG 780
Qy 1890 TGGTTTCAATAGTCTCTCTCAGCAGCATCTCTGATTTACCTTCGGAACAGTGTGATGC 1949
Db 781 CGGCTTCAACAGCTCTCTCTCAGTGAAGCTTCTTGTCTTACCTTGGAGCAGTCCGATGC 840
Qy 1950 GCATCGGAGTGAACCGCTGTTGTTAGCAGCAGGACCTGGTATTTGGTGTGTGCGAACAC 2009
Db 841 ACACCGAGTGAAGCTGTTGTTGTTGGCAGCAGGACCTGGTATCCGGTGTCTGTGCGAAC 900
Qy 2010 AGGTCGTCTCAGTGTATCTGTTGGCGCTGCGCAACTGATCAACAGAGAAAAGTATAA 2069
Db 2010 AGGTCGTCTCAGTGTATCTGTTGGCGCTGCGCAACTGATCAACAGAGAAAAGTATAA 2069
```

Db 901 ACAGTCGTCTCGATGTACCTCTCGGAGTTGGCAAAGTGAAGAAAGTAA 960
QY 2070 ATCAGATGTTTTCAAAAGAACTCTGACCATGACAGATGTGACCAACACAGATTG 2129
Db 961 ATCAGAGTGTCTTAAAGAAACCTTGACCATGACAGATGTGACCAACACAGATTG 1020
QY 2130 TTACAGGTGTACAGCCAAACCAATGACTGCCACTGGTCAATGACCATTTGTCCCCAG 2189
Db 1021 TTACAGTGTACAGCCAAATACCAATGACTGCCACTGGTCAATGACCATTTGTCCCTGT 1080
QY 2190 GAACCAAGCTGTCTAGAGCCAGATCTCCATTTTATGATGAGAAATGCCCCAAGGA 2249
Db 1081 GAACCAAGCTGTACAGAAAGCCAGATCTCCATTTGCCAATGAGATGTGCCCAAGGA 1140
QY 2250 TAACCCCATGTACTACTGTAAACAAAGAACCCAGCTCAGAGAGCTGGCCCTGCAGCA 2309
Db 1141 TAACCCCATGTACTACTGCATTAAGAAACCCAGCTCAGAGAGCTGGCCCTGCAGCA 1200
QY 2310 CTGCCAGTGGAGCCCGGAATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCAT 2369
Db 1201 CTGCCAGTGGAGCCCGGAATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCA 1260
QY 2370 TGCTCGCATTTGGTTGGAACTCATGTTTGAATTAATTAATTAATTAATTAATTAATTA 2429
Db 1261 TGCTCGCATTTGGTTGGAACTCATGTTTGAATTAATTAATTAATTAATTAATTAATTA 1320
QY 2430 CAATGCTAAATTTTCTGTAGAACCAAAATGCCCTTTGGCTTCTTTACAAACCAAG 2489
Db 1321 CAATGCTAAATTTTCTGTAGAACCAAAATGCCCTTTGGCTTCTTTACAAACCAAG 1380
QY 2490 GAAGTGAATTTGCTTAAGCAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTA 2549
Db 1381 GAAGTGAATTTGCTTAAGCAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTA 1440
QY 2550 GCTCACTTTAAACCCATGGTGGCTTCGGAAGATCAATGTGCTTCTGCTGCTGGGA 2609
Db 1441 GCTCACTTTAAACCCATGGTGGCTTCGGAAGATCAATGTGCTTCTGCTGCTGGGA 1500
QY 2610 AGATATGCTCCCATTAATAATAGTTTACTACAGTGAATGGCTGTAGCCCAAGTATGC 2669
Db 1501 GGATATGCTCCCATTAATAATAGTTTACTACAGTGAATGGCTGTAGCCCAAGTATGC 1560
QY 2670 TGATTTCTGTGAATTTTATCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAT 2729
Db 1561 TGCTTTCTGTGGATTTCTAGAGCTTAGTACTCGGGATTAAGGCTGCAACCTGCAT 1620
QY 2730 CAACCACTCAATAGTGTGTGTAAAGGCTGTCAAAACCAACAGTGTGAAGAGTGC 2789
Db 1621 CAACCACTCAATAGTGTGTGTAAAGGCTGTCAAAACCAACAGTGTGAAGAGTGC 1680
QY 2790 GACACATGTGCTTGAGGACAGCATGTGGAGATGACACAGCGGAGCTCTGAGTGCAT 2849
Db 1681 GACACATGTGCTTGAGGACAGCATGTGGAGATGACACAGCGGAGCTCTGAGTGCAT 1740
QY 2850 GTGGTGCAGCAACATGAAGCAGTGTGTGAGTCTCAATGCTTATGTGGCTCTCTTCCCTTT 2909
Db 1741 GTGGTGCAGCAACATGAAGCAGTGTGTGAGTCTCAATGCTTATGTGGCTCTCTTCCCTTT 1800
QY 2910 TGGCCAGTGTATGNAATGGTATACATGAGCAGCACTGCCCTGAAATTTGTTGAGCTTA 2969
Db 1801 TGGCCAGTGTATGNAATGGTATACATGAGCAGCACTGCCCTGAAATTTGCTGTGCTTA 1860
QY 2970 CTGTACCTGTAGTCAATTTCTTGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAA 3029
Db 1861 TTGTACCTGTGAGCAATTTCTTGGAGCAGCAGGCTGTGGTGTGTACTGATCCAGCAA 1920
QY 3030 TACTGGCAAGGAAATGCAATAGAGGTTTCTATTAAGGACCAAGTGAAGTGCCTTGC 3089
Db 1921 TACTGGCAAGGAAATGCAATAGAGGTTTCTATTAAGGACCAAGTGAAGTGCCTTGC 1980
QY 3090 AGCCCTTACAGGAAATTTCTATCAGACGCTCTGCTCAATTTCCAGCATGTGTCTAGAGGA 3149
Db 1981 GGCCTCTCAGGAAATGTTATCCACAGCCCTTCTGAACTCCAGCATGTGTCTAGAGGA 2040

QY 3150 CAGCAGATACAACTGGTCTTTTCACTGTCCAGCTTGCATATGCAACGCGCCACAGTAA 3209
Db 2041 CAGCAGATACAACTGGTCTTTTCACTGTCCAGCTTGCATATGCAACGCGCCACAGTAA 2100
QY 3210 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACTGACCAACAGGCAAGCACTGCGA 3269
Db 2101 ATGCATCAACAGAGTGTCTGTGAGAAGTGTGAGAACTGACCAACGCGCCAGCACTGCGA 2160
QY 3270 GACCTGCATATCTGGCTTTCTAGCGGTATCCCAACCAATGAGGAGAAATGTTCAGCCATGCAA 3329
Db 2161 GACCTGCATATCTGGCTTTCTAGCGGTATCCCAACCAATGAGGAGAAATGTTCAGCCATGCAA 2220
QY 3330 GTGCAATGGGCAAGCGCTCTGTGCAACCAACAGGCAAGTGTCTGCAACCAACCA 3389
Db 2221 GTGCAATGGGCAAGCGCTCTGTGCAACCAACAGGCAAGTGTCTGCAACCAACCA 2280
QY 3390 GGGCTCAAGGGGAGCGAGTCCAGCTATGTAGGTAGAAATCGATACCAAGAAACCC 3449
Db 2281 AGGTGTCAAGGGGAGCGAGTCCAGCTATGTAGGTAGAAATCGATACCAAGAAATCC 2340
QY 3450 TCTCAGAGGAACATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3509
Db 2341 TCTCAGAGGAACATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
QY 3510 CCAGGAAGATGATCGCTTTTACACAGCTATCAATTTTGGCTTACTCTGACCAACAAA 3569
Db 2401 CCAGGAAGATGATCGCTTTTACACAGCTATCAATTTTGGCTTACTCTGACCAACAAA 2460
QY 3570 CAGGATTTGGACATGTTTCAATCAATGCTCCAAAGAAATTAACCTCAACATCACTGGGC 3629
Db 2461 CAGGATTTGGACATGTTTCAATCAATGCTCCAAAGAAATTAACCTCAACATCACTGGGC 2520
QY 3630 TGCCAGTTTCTCAGCTGGAAACCCAGCTGAGAGAGATGCTGTGTTTCAAAACCAA 3689
Db 2521 CACCAGTTTCCAGCCGGAACCCAGCTGAGAGAGATGCTGTGTTTCAAAACCAA 2580
QY 3690 CATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTGCAACCAACCAA 3749
Db 2581 CATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTGCAACCAACCAA 2640
QY 3750 TATCAGTTTCTTGTGTTTGTGATGATTAATTTCACTGGCCCAACAAATTCAGATGCTTT 3809
Db 2641 CATCAGTTTCTTGTGTTTGTGATGATTAATTTCACTGGCCCAACAAATTCAGATGCTTT 2700
QY 3810 CTCTCAGCAGCAATTTTATGGAACCTGGTACAGTTCTTTGCTGACTTTTTCAGTTGTTT 3869
Db 2701 CTCTCAGCAGCAATTTTATGGAACCTGGTACAGTTCTTTGCTGACTTTTTCAGTTGTTT 2760
QY 3870 CTTCTCTTTGCTTCTGGTGGCTGTGTTTGGAGATCAACAAAGTTGTTGGGCTC 3929
Db 2761 CTTCTCTCTGCTTCTGGTGGCTGTGTTTGGAGATCAACAAAGTTGTTGGGCTC 2820
QY 3930 CAGACGTAGAGCAACTTTCTTCGAGAGATGCAACAGATGGCCGCTCCCTTTGCTC 3989
Db 2821 CAGGCGGAGAGCAACTTTCTTCGAGAGATGCAACAGATGGCCGCTCCCTTTGCTC 2880
QY 3990 TGTAAATGTGCTTGGAAACAGATGAGAGGCTCTGATCTTATTTGGGGGAGTATAA 4049
Db 2881 TGTAAACGTGTGCTTGGAAACAGATGAGAGGCTCTGATCTTATTTGGGGGAGTATAA 2940
QY 4050 GACTGTTTCCAAACCATTTGCACTGAGCGTGTGTTTGGCAACAAAGCGCTGCTCTC 4109
Db 2941 GACCGTTCCTAAGCCCATTTGCCCTGGAGGCTGCTTTGGTAACAAGCGCAGTCTCTC 3000
QY 4110 TGTGTTGTGAGGCTTCCCTGAGGCTGGTGGCATTCCTCTCTGCTGGCAGTCAAGTCT 4169
Db 3001 TGTATTTGCTGAGGCTTCCCTGAGGCTGGAGGAAATCCCTCTCTGCTGCTGAGTCT 3060
QY 4170 TCGTGTGGCCAGCGCTTGTGCAATTTCTCAGCAGATGCTGATGTAACAGGAGAA 4229
Db 3061 CGCTGTGGCCAGTGGCTTGTGCAATTTCTCAGCAGATGCTGATGTAACAGGAGAA 3120

QY 1683 AGTGAGTGGAAACCATGCTGGTGTGGGGGAACACACACAATGACATCTATGAGCCA 1742
DB |||||
QY 1680 AGTGAGTGGAAACCATGCTGGTGTGGGGGAACACACACAATGACATCTATGAGCCA 1739
DB |||||
QY 1743 TGGCGCAAAATGCTCTCTCTCGAATTTTCATGGCTATGACATTTGCTGTGACCGCTGGTC 1802
DB |||||
QY 1740 TGGCGCAAAATGCTCTCTCTCGAATTTTCATGGCTATGACATTTGCTGTGACCGCTGGTC 1799
DB |||||
QY 1803 AGTGCTTCCGAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862
DB |||||
QY 1800 AGTGCTTCCGAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1859
DB |||||
QY 1863 ACACAAACAGCACCATGATGTGTGGTGGTTCATATGATCTCTCTCCAGCGCATCTCT 1922
DB |||||
QY 1860 ACACAAACAGCACCATGATGTGTGGTGGTTCATATGATCTCTCTCCAGCGCATCTCT 1919
DB |||||
QY 1923 GGTATTCACTCGGAACAGTGTGCGCATCGGAGTGAAGCGGCTTTGTTAGCAGCAGG 1982
DB |||||
QY 1920 GGTATTCACTCGGAACAGTGTGCGCATCGGAGTGAAGCGGCTTTGTTAGCAGCAGG 1979
DB |||||
QY 1983 AACTGATATTCGGTGTGGTGGAAACACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGC 2042
DB |||||
QY 1980 AACTGATATTCGGTGTGGTGGAAACACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGC 2039
DB |||||
QY 2043 AACTGATGAACAGGAAGAAAGTTAAATTCAGAAATGTTTTTCCAAAGAACTCTTGACCA 2102
DB |||||
QY 2040 AACTGATGAACAGGAAGAAAGTTAAATTCAGAAATGTTTTTCCAAAGAACTCTTGACCA 2099
DB |||||
QY 2103 TCACAGATGTGACAGGACACACAGATTTGTACAGTGTACAGGTCACAGGTCACAGGTCACCA 2162
DB |||||
QY 2100 TCACAGATGTGACAGGACACACAGATTTGTACAGTGTACAGGTCACAGGTCACAGGTCACCA 2159
DB |||||
QY 2163 CTGGTGCATGACCATGCTGTGCTCCAGGAACACACAGTGTCTCAGAGGCCAGATCTCCAT 2222
DB |||||
QY 2160 CTGGTGCATGACCATGCTGTGCTCCAGGAACACACAGTGTCTCAGAGGCCAGATCTCCAT 2219
DB |||||
QY 2223 TTTTAGGTATGAGATTTGCCCAAGGATAACCCATGATCTGTTAAACAGAGACCAAG 2282
DB |||||
QY 2220 TTTTAGGTATGAGATTTGCCCAAGGATAACCCATGATCTGTTAAACAGAGACCAAG 2279
DB |||||
QY 2283 CTGACAGGAGCTGTGCTCCAGGACCAAGCTGCTGAGGAGCCCGGATCAGGAGTGCAT 2342
DB |||||
QY 2280 CTGACAGGAGCTGTGCTCCAGGACCAAGCTGCTGAGGAGCCCGGATCAGGAGTGCAT 2339
DB |||||
QY 2343 TGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGG 2387
DB |||||
QY 2340 TGCCCTGCCGAGTGGCTTGCAGGGTCTATCTTGGTGTGTGGG 2384
DB |||||

RESULT 14

AR164815
LOCUS AR164815 6370 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815
VERSION AR164815.1 GI:16238024
KEYWORDS
SOURCE Unknown.

ORGANISM

Unclassified.

1 (bases 1 to 6370)

Moore,K. and Nagle,D.Lynn.

Methods and compositions for the diagnosis and treatment of body

weight disorders, including obesity

Patent: US 6274339-A 12-14-AUG-2001;

location/Qualifiers

1. 6370

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 43.2%; Score 1853; DB 6; Length 6370;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;

QY 2257 ATGTACTACTGTAAACAAGAACAGCTGTCAGGAGCTGTGCCCTGGACACAGAACTGCCAG 2316
DB |||||
QY 2317 TGGAGACCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAAAATATCTGTGGCATTTGCTGG 2376
DB |||||
QY 61 TGGAGACCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAAAATATCTGTGGCATTTGCTGG 120
DB |||||
QY 2377 CATTTGGTTGGAAACTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATGACAATGCT 2436
DB |||||
QY 121 CATTTGGTTGGAAACTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATGACAATGCT 180
DB |||||
QY 2437 AAATTGTTCTGTAGGAACACAAATGCGCTTTTGGCTTCTTTACAAACCCAGAAAGGTA 2496
DB |||||
QY 181 AAATTGTTCTGTAGGAACACAAATGCGCTTTTGGCTTCTTTACAAACCCAGAAAGGTA 240
DB |||||
QY 2497 GAATTTGCTTAAAGCAGCTGCGAAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACC 2556
DB |||||
QY 241 GAATTTGCTTAAAGCAGCTGCGAAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACC 300
DB |||||
QY 2557 TTAACCCCATGGGTGCGGCTT---CGGAAGATCAATGTCTCTACTGG---TGCTGG 2608
DB |||||
QY 301 TTAACCCCATGGGTGCGGCTTTCGGGAAGGTYCAATGTCTCTACTKGGTGTGGGGAAG 360
DB |||||
QY 2609 AAGATATGTCCTCATTTACAAATAG---TTTACTACAGTGGATGCC---GTCTGAGCCAG 2663
DB |||||
QY 361 GATATGTCCTCATTTTACAAATAGTTTACTACAGTGGGATGSCCGTCTTGAGGCCAG 420
DB |||||
QY 2664 TGATGCT---GGATTCCTGTGGAATTTT---ATCAGAACCCAGTACTCGGGGACTGAAG 2716
DB |||||
QY 421 TGTGCTTGGATTTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGACTGAAG 480
DB |||||
QY 2717 CTGCAACCTGCAT---CAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAACACAG 2774
DB |||||
QY 481 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAACACAG 540
DB |||||
QY 2775 TGCTAA---GCAGTGGCCGACACCATGTGCTTCAGACACAGCATGTGGAGATTCACACAG 2833
DB |||||
QY 541 TGCTAAGGAGTGGCCGACACCATGTGCTTCAGACACAGCATGTGGAGATTCACACAG 600
DB |||||
QY 2834 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGAATTCCTCAATGCTATG 2893
DB |||||
QY 601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGAATTCCTCAATGCTATG 660
DB |||||
QY 2894 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGCACTGCCCCCTG 2953
DB |||||
QY 661 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGCACTGCCCCCTG 720
DB |||||
QY 2954 AAAATTGTTACGGCTACTGTACTGTAGTCTATGCTTGGAGCAACAGGCTGTGCTGCT 3013
DB |||||
QY 721 AAAATTGTTACGGCTACTGTACTGTAGTCTATGCTTGGAGCAACAGGCTGTGCTGCT 780
DB |||||
QY 3014 GTACTGATCCCGCAATACTGCGAAGGGAATGCAATAGAGGTTCTCTATTAAGACACAG 3073
DB |||||
QY 781 GTACTGATCCCGCAATACTGCGAAGGGAATGCAATAGAGGTTCTCTATTAAGACACAG 840
DB |||||
QY 3074 TGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCCA 3133
DB |||||
QY 841 TGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCCA 900
DB |||||
QY 3134 GCATGTGCTAGAGACAGCAGATACAACTGGTCTTTTCAATTCATGCTCCAGCTGCCAAT 3193
DB |||||
QY 901 GCATGTGCTAGAGACAGCAGATACAACTGGTCTTTTCAATTCATGCTCCAGCTGCCAAT 960
DB |||||
QY 3194 GCAACGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGNACCTGACCA 3253
DB |||||
QY 961 GCAACGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGNACCTGACCA 1020
DB |||||
QY 3254 CAGGCAACACTGGAGACCTGCAATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGA 3313
DB |||||
QY 1021 CAGGCAACACTGGAGACCTGCAATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGA 1080
DB |||||

QY 2954 AAAATTGTTCAAGGCTACTGTACCTGTAGTCAATTCCTTGGAGCAACCAAGGCTGTGGCTGGT 3013
Db AAAATTGTTCAAGGCTACTGTACCTGTAGTCAATTCCTTGGAGCAACCAAGGCTGTGGCTGGT 420
QY 3014 GTACTGATCCAGCAATACTGGCAAGGGAAATCATAGAGGTTCTATAAAGGACCAG 3073
Db GTACTGATCCAGCAATACTGGCAAGGGAAATCATAGAGGTTCTATAAAGGACCAG 480
QY 3074 TGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTCTCAATTCGA 3133
Db TGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTCTCAATTCGA 540
QY 3134 GCATGTCTAGAGACAGCAGATACAACTGGTCTTTTCATCTGCTCCAGCTTGCAT 3193
Db GCATGTCTAGAGACAGCAGATACAACTGGTCTTTTCATCTGCTCCAGCTTGCAT 600
QY 3194 GCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCA 3253
Db GCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCA 660
QY 3254 CAGGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTGTATCCCAATGGAGGGA 3313
Db CAGGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTGTATCCCAATGGAGGGA 720
QY 3314 AATGTACGCCATGCAAGTGAATGGGCAACGCTCTGTGCAACACCAACACCGGCAAGT 3373
Db AATGTACGCCATGCAAGTGAATGGGCAACGCTCTGTGCAACACCAACACCGGCAAGT 780
QY 3374 GCTTCTGCAACCAACGAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAAATC 3433
Db GCTTCTGCAACCAACGAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAAATC 840
QY 3434 GATACCAAGGAAACCTCTCAGAGGACATGTTATTAATCTCTTATTAATGACTATCAGT 3493
Db GATACCAAGGAAACCTCTCAGAGGACATGTTATTAATCTCTTATTAATGACTATCAGT 900
QY 3494 TCACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3553
Db TCACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 960
QY 3554 CTCCTGACGAAACAAACAGGGAATTTGACATGTTCAATGCTCCAGAAATTTCAACC 3613
Db CTCCTGACGAAACAAACAGGGAATTTGACATGTTCAATGCTCCAGAAATTTCAACC 1020
QY 3614 TCAACATCAGCTGGCTGCCAGTTTCTCAGTGGACCCAGGCTGGAGAGAGATGCTG 3673
Db TCAACATCAGCTGGCTGCCAGTTTCTCAGTGGACCCAGGCTGGAGAGAGATGCTG 1080
QY 3674 TTGTTTCAAAAAACCAATTAAGAGGTACAAAGATAGTTTCTCTAATGAGAGATTTGAT 3733
Db TTGTTTCAAAAAACCAATTAAGAGGTACAAAGATAGTTTCTCTAATGAGAGATTTGAT 1140
QY 3734 TTGCAACCAACCAAAATCACTTTCTTTGTTATGTCAGTAATTTCACTGGCCCATCA 3793
Db TTGCAACCAACCAAAATCACTTTCTTTGTTATGTCAGTAATTTCACTGGCCCATCA 1200
QY 3794 AAATTCAGATTCCTCTCTCAGCAGCAAGCAATTTATGGACCTGGTACAGTCTTCGTGA 3853
Db AAATTCAGATTCCTCTCTCAGCAGCAAGCAATTTATGGACCTGGTACAGTCTTCGTGA 1260
QY 3854 CTTTCTTCAGTGTGTTCT 3913
Db CTTTCTTCAGTGTGTTCT 1320
QY 3914 AAAGTTGTTGGGCTCCAGACGTAGAGGCAACTTCTCTGAGAGATGCAACAGATGGCCA 3973
Db AAAGTTGTTGGGCTCCAGACGTAGAGGCAACTTCTCTGAGAGATGCAACAGATGGCCA 1380
QY 3974 GCGTCCCTTTCGCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCCTCTCTGATCTTA 4033
Db GCGTCCCTTTCGCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCCTCTCTGATCTTA 1440
QY 4034 TTGGGGGAGTATAAAGACTGTTCCCAAAACCCATTCGACTGGAGCGTGTGTTGGCAACA 4093

Db 1441 TTGGGGGAGTATAAAGACTGTTCCCAAAACCCATTCGACTGGAGCGTGTGTTGGCAACA 1500
QY 4094 AAGCCGCTGTCCCT 4153
Db AAGCCGCTGTCCCT 1560
QY 4154 CTGGGCAAGTCAGGCT 4213
Db CTGGGCAAGTCAGGCT 1620
QY 4214 TAGTGTACAGGAGAAAGTCAGGAGCCGTGAGAAACCCGGAAGCAGAGCCCTCTGACAGC 4273
Db TAGTGTACAGGAGAAAGTCAGGAGCCGTGAGAAACCCGGAAGCAGAGCCCTCTGACAGC 1680
QY 4274 CTGGGACCTGCACTGGA 4290
Db CTGGGACCTGCACTGGA 1697

Search completed: March 5, 2004, 05:11:31
Job time : 10814.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:59:34 ; Search time 6797.26 Seconds

(without alignments)
18847.119 Million cell updates/sec

Title: US-09-787-097-13

Perfect score: 4290

Sequence: 1 atgtgtggcgcagcgccgc.....agcctggaccgcgcatctga 4290

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3429.2	79.9	3976	29	AV418587 Homo sapi
2	2923.6	68.1	3976	29	AV418589 Mus muscu
3	2864.4	66.8	3671	29	AV418588 Pan trogl
4	925.6	21.6	1201	13	BX440935 BX440935

5	775	18.1	964	13	BU506373
6	758.8	17.7	933	12	BG678679
7	701.2	15.3	840	14	CB520374
8	681	15.9	681	12	BM783739
9	680.4	15.9	683	12	BM783788
10	680	15.9	680	12	BM783258
11	680	15.9	680	12	BM783773
12	677	15.8	677	12	BM783718
13	674.4	15.7	677	12	BM783821
14	673	15.7	673	12	BM783685
15	663.8	15.5	792	14	CD644584
16	652.2	15.2	962	13	EX431593
17	636.8	14.8	674	12	BG563729
18	636.4	14.8	770	13	BY761276
19	636	14.8	777	14	CF745736
20	632.4	14.7	706	12	BI089332
21	632	14.7	632	12	BM773003
22	632	14.7	632	12	BM783264
23	626.6	14.6	953	10	BF313142
24	623	14.5	917	10	BF672370
25	618	14.4	633	10	B8894518
26	610.6	14.2	722	13	BX669806
27	608	14.2	735	13	BY761390
28	607.8	14.2	991	10	B8880905
29	597.6	13.9	940	13	BQ956223
30	589.6	13.7	743	14	CB519323
31	585.2	13.6	716	14	CF745901
32	580.2	13.5	714	10	BE650220
33	580	13.5	591	12	BI850289
34	568.2	13.2	653	14	CB465934
35	554.2	12.9	663	14	CP745835
36	548	12.8	902	13	BU151781
37	537.4	12.5	644	14	CB425680
38	535.8	12.5	942	10	BF316716
39	530.6	12.4	641	10	B8664616
40	524.2	12.2	626	10	BE980188
41	521	12.1	600	14	CB432517
42	517.6	12.1	595	10	B853681
43	516.6	12.0	598	10	B853006
44	513.2	12.0	544	10	B853962
45	502.4	11.7	567	10	BF854487

ALIGNMENTS

RESULT 1	AV418587	3976 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,				
DEFINITION	Genomic survey sequence.				
ACCESSION	AY418587				
VERSION	AY418587.1	GI:39774547			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 3976) Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 3976)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..3976

organism="Homo sapiens"

mol_type="genomic DNA"

db_xref="taxon:9606"

gene <1..>3976

locus_tag="HCM6607"

ORIGIN

Query Match 79.9%; Score 3429.2; DB 29; Length 3976;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 3431; Conservative 0; Mismatches 545; Indels 0; Gaps 0;

QY 315 TGTCAACGCGCTGCTGCAACCTGGCACCGCCAGTGGCTCTGCCCGCGCGGTGGT 374

Db 1 TGTCAACGCGCTGCTGCAACCTGGCACCGCCAGTGGCTCTGCCCGCGCGGTGGT 60

QY 375 GGGCAGCAATCCAGCACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGTGGT 434

Db 61 GGGCAGCAATCCAGCACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGTGGT 120

QY 435 GACAGATGACCTGGAAATATATAAATAAACAAGAGTGGACGTGGCTCATTTGAAGACA 494

Db 121 NNN 180

QY 495 GCCAATAGAAATATGAGACTTCGTTTCAATTCATTTGCTACAGAGTGTAGTGGACCA 554

Db 181 GCCAATAGAAATATGAGACTTCGTTTCAATTCATTTGCTACAGAGTGTAGTGGACCA 240

QY 555 TTATATGTTTATGATGGGACTCAATTTATGCACCGCTAGTCTGCTCATTTAGTGGCCT 614

Db 241 TTATATGTTTATGATGGGACTCAATTTATGCACCGCTAGTCTGCTCATTTAGTGGCCT 300

QY 615 CATTGTTCTTGAGAGATGCCATGAGACTGTCCCTGAGGTGTTGGCCATCAGTTA 674

Db 301 NNN 360

QY 675 TGCCTTGCTGCTATTTTATGATGCTGTCTATAATTTGACTGGAATTAATTACTTA 734

Db 361 NNN 420

QY 735 CAGTTTTCATATGTCCTCAATAACTGCTCAGCCGAGGAGGTGAAGATCAGTAATAG 794

Db 421 NNN 480

QY 795 CAGCGAACTCTTCAATGTGAATGTTCTGAAATCTGAAAGGTGAAGCATGTGACATTC 854

Db 481 CAGCGAACTCTTCAATGTGAATGTTCTGAAATCTGAAAGGTGAAGCATGTGACATTC 540

QY 855 TCACGTACAGCAACTGTGGTTTCTCATPCGAGGCATCTGCAATTCAAAGTATGTCAG 914

Db 541 TCACGTACAGCAACTGTGGTTTCTCATPCGAGGCATCTGCAATTCAAAGTATGTCAG 600

QY 915 AGGATGCTCCTGCTTCACAGCTGGCAGGTCTGAGGTGTTGAGTCTGCTGACCTAA 974

Db 601 AGGATGCTCCTGCTTCACAGCTGGCAGGTCTGAGGTGTTGAGTCTGCTGACCTAA 660

QY 975 CCAGTCAATTTGAGCTCGAGAGGAATATCTAACTTAAAGCTCCCGAGCATCTCATAA 1034

Db 661 CCAGTCAATTTGAGCTCGAGAGGAATATCTAACTTAAAGCTCCCGAGCATCTCATAA 720

QY 1035 AGCTGCTGCTCAATGGAACATATGTTGGTGTGAGGATATATGTTCAACACTCAGA 1094

Db 721 AGCTGCTGCTCAATGGAACATATGTTGGTGTGAGGATATATGTTCAACACTCAGA 780

QY 1095 TTATAACATGGTTCTAGCGTATGACCTTGTCTAGGGAGTGGCTTCCACTAACCGTTC 1154

Db 781 TTATAACATGGTTCTAGCGTATGACCTTGTCTAGGGAGTGGCTTCCACTAACCGTTC 840

QY 1155 TGTCAACAAATGTGGTTGTAGATATGTCATTCCTTTGGCATTTATACAAGGATAAAATTTA 1214

Db 841 TGTCAACAAATGTGGTTGTAGATATGTCATTCCTTTGGCATTTATACAAGGNNNNNNNN 900

QY 1215 CATGTATGGAGGAAAAATTCATCAACTGGAAATGTGACCAATGAGTTCAGAGTTTTC 1274

Db 901 NNN 960

QY 1275 CATTCAATAGACTCATGGTGTGTTGACCCCTAAGGCAAGGACGATATGCGAGTGT 1334

Db 961 NNN 1020

QY 1335 TGGGCACTCTGCACACATTTGTACACTGAAGAATGCCGAGTGGTCACTGCTGTCATCTT 1394

Db 1021 NNN 1080

QY 1395 TGTCTACTGCCCTCTCTATGGATATATAAGCAATGTGCAAGGAATATGATTTGGATAAG 1454

Db 1081 NNN 1140

QY 1455 CACATCGAGTATATTACACACCCAGGTCCTTGTGCAAGGGGTACGGCCATAGCAG 1514

Db 1141 CACATCGAGTATATTACACACCCAGGTCCTTGTGCAAGGGGTACGGCCATAGCAG 1200

QY 1515 TGTTTACGACCATAGGACCAAGGCGCTTATACGTTTATGTTGGCTTACAAGGCTTTCAGTGC 1574

Db 1201 TGTTTACGACCATAGGACCAAGGCGCTTATACGTTTATGTTGGCTTACAAGGCTTTCAGTGC 1260

QY 1575 CAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGCGATACCCAGATGAGAC 1634

Db 1261 CAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGCGATACCCAGATGAGAC 1320

QY 1635 CATTTCTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGATAGTGAAGTGAAC 1694

Db 1321 CATTTCTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGATAGTGAAGTGAAC 1380

QY 1695 CATGCTGTTGTTGGGGGAAACACACAAATGACACATCTATAGCCCATGGCGCAATATG 1754

Db 1381 CATGCTGTTGTTGGGGGAAACACACAAATGACACATCTATAGCCCATGGCGCAATATG 1440

QY 1755 CTTCTCTTCAGATTTTCAATGCTTATGACCTGCTGTCACCGCTGCTGAGTCTGCCAG 1814

Db 1441 CTTCTCTTCAGATTTTCAATGCTTATGACCTGCTGTCACCGCTGCTGAGTCTGCCAG 1500

QY 1815 ACTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACACAGCAC 1874

Db 1501 NNN 1560

QY 1875 CATGTATGTTGCGTGGTTCATAGTCTCCTCTCAGCGACATCTGTTATTCACCTC 1934

Db 1561 CATGTATGTTGCGTGGTTCATAGTCTCCTCTCAGCGACATCTGTTATTCACCTC 1620

QY 1935 GGAAAGTGTGATGCGCATCGAGTGAAGCGCTTGTTTAGCAGCAGACCTGTTATTCG 1994

Db 1621 GGAAAGTGTGATGCGCATCGAGTGAAGCGCTTGTTTAGCAGCAGACCTGTTATTCG 1680

QY 1995 GGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAACA 2054

Db 1681 GGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAACA 1740

QY 2055 AGAAGAAAGTTAAATCAGAAATGTTTTTCAAAAGAACTCTTCACCATGACAGATGTGA 2114

Db 1741 AGAAGAAAGTTAAATCAGAAATGTTTTTCAAAAGAACTCTTCACCATGACAGATGTGA 1800

QY 2115 CCAGACACAGATTTGTTACGCTGTACGCGAAACACCAATGATGCGCATGCGTGAATGA 2174

Db 1801 CCAGACACAGATTTGTTACGCTGTACGCGAAACACCAATGATGCGCATGCGTGAATGA 1860

QY 2175 CCATTTGTCGCCAGGAACCAACAGCTGCTCAGAGGCGCAGATCTCCATTTTATGATGA 2234

Db 1861 CCATTTGTCGCCAGGAACCAACAGCTGCTCAGAGGCGCAGATCTCCATTTTATGATGA 1920

QY	2235	GAATTGGCCCAAGGATTAACCCCATGTACTACTGTATCAAGAAAGACAGCTGCGAGGAGCTG	2294
Db	1921	GAATTGGCCCAAGGATTAACCCCATGTACTACTGTATCAAGAAAGACAGCTGCGAGGAGCTG	1980
QY	2295	TGCCCTGGACAGCAACTGCGAGTGGAGCCCGGGAATCAGAGTGCAATTCGCCCTGCCCCGA	2354
Db	1981	TGCCCTGGACAGCAACTGCGAGTGGAGCCCGGGAATCAGAGTGCAATTCGCCCTGCCCCGA	2040
QY	2355	AAATATCTGTGGCAATGCGCTGGCAATTTGGTGTGAAACTCATGTTTGAATAATTAAGTATGTC	2414
Db	2041	AAATATCTGTGGCAATGCGCTGGCAATTTGGTGTGAAACTCATGTTTGAATAATTAAGTATGTC	2100
QY	2415	CAGGAGAAATATCAGCAATGCTAAATTTCTGTAGGAACCAACAATCCCTTTTGGCTTC	2474
Db	2101	CAGGAGAAATATCAGCAATGCTAAATTTCTGTAGGAACCAACAATCCCTTTTGGCTTC	2160
QY	2475	TCATTACACCCAGAGAGTAGAATTTGTCCTTTAAGCAGCTGCGAATAATGCAAGTCAATC	2534
Db	2161	TCATTACACCCAGAGAGTAGAATTTGTCCTTTAAGCAGCTGCGAATAATGCAAGTCAATC	2220
QY	2535	TCAGAGCATGTCGAAGCTCACTTAACCCCAAGGTTGGGCTTCGGAAGATCAATGTGTC	2594
Db	2221	TCAGAGCATGTCGAAGCTCACTTAACCCCAAGGTTGGGCTTCGGAAGATCAATGTGTC	2280
QY	2595	CTACTGGTCTGGGAAGATATGTCCTTACCAATTTACAAATAGTTTACTACAGTGGATGCCGTC	2654
Db	2281	CTACTGGTCTGGGAAGATATGTCCTTACCAATTTACAAATAGTTTACTACAGTGGATGCCGTC	2340
QY	2655	TCAGCCAGTGAATGCTGTGGAATTTTATCAGAACCCAGTACTGGGAGCTGAA	2714
Db	2341	TCAGCCAGTGAATGCTGTGGAATTTTATCAGAACCCAGTACTGGGAGCTGAA	2400
QY	2715	GGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCAACAG	2774
Db	2401	GGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCAACAG	2460
QY	2775	TGCTAAGCAGTCCCGGACCACTATGCTTCCAGACAGCATGTGGAGANTGCCACAGGG	2834
Db	2461	TGCTAAGCAGTCCCGGACCACTATGCTTCCAGACAGCATGTGGAGANTGCCACAGGG	2520
QY	2835	CAGCTCTCAGTGCATGTGGTCAGCAACATGAGCAGTGTGACCTCCATGCTATGT	2894
Db	2521	CAGCTCTCAGTGCATGTGGTCAGCAACATGAGCAGTGTGACCTCCATGCTATGT	2580
QY	2895	GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCCTGA	2954
Db	2581	GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCCTGA	2640
QY	2955	AAATGTTACAGCTACTGTAACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGCTG	3014
Db	2641	AAATGTTACAGCTACTGTAACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGCTG	2700
QY	3015	TACTGATCCAGCAATACCTGGCAAGGAAATGCAATAGAGGTTTCTTATAAGGACCACT	3074
Db	2701	TACTGATCCAGCAATACCTGGCAAGGAAATGCAATAGAGGTTTCTTATAAGGACCACT	2760
QY	3075	GAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG	3134
Db	2761	GAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG	2820
QY	3135	CATGTGCTTAGAGCACAGATACAACTGGTCTTTCTACTCTCCAGCTTGGCAATG	3194
Db	2821	CATGTGCTTAGAGCACAGATACAACTGGTCTTTCTACTCTCCAGCTTGGCAATG	2880
QY	3195	CAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCAAC	3254
Db	2881	CAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCAAC	2940
QY	3255	AGCAAGACCTGGAGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGAA	3314
Db	2941	AGCAAGACCTGGAGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGAA	3000
QY	3315	ATGTACGCCATGCAAGTGAATGGGCAACCGCTCTCTGTGCAACACCAACAGCGGCAAGTG	3374

Db	3001	ATGTACGCCATGCAAGTGCATGGGCAACCGCTCTGTGCAACACCAACACGGGCAAGTG	3060
QY	3375	CTTCTGCAACCAAGAGGGCTCAAGGGGACAGGTGCGAGCTATGTGAGTGAAGAAATCG	3434
Db	3061	CTTCTGCAACCAAGAGGGCTCAAGGGGACAGGTGCGAGCTATGTGAGTGAAGAAATCG	3120
QY	3435	ATACCAAGGAAACCTCTCAGAGGAACATGTTATTATATACTCTTCTTATTGACTATCAGTT	3494
Db	3121	ATACCAAGGAAACCTCTCAGAGGAACATGTTATTATATACTCTTCTTATTGACTATCAGTT	3180
QY	3495	CACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGCTAC	3554
Db	3181	CACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGCTAC	3240
QY	3555	TCCTGACGAACAAACACAGGATTTGGACATGTTCAATGCTCCAGAAATTTCAACCT	3614
Db	3241	TCCTGACGAACAAACACAGGATTTGGACATGTTCAATGCTCCAGAAATTTCAACCT	3300
QY	3615	CAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTCT	3674
Db	3301	CAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTCT	3360
QY	3675	TGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATTT	3734
Db	3361	TGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATTT	3420
QY	3735	TGCCAACCAACCAACATTAATCACTTTCTTTTATGTAGTAAATTTTCACTGGCCCACTAA	3794
Db	3421	TGCCAACCAACCAACATTAATCACTTTCTTTTATGTAGTAAATTTTCACTGGCCCACTAA	3480
QY	3795	AATTCAGATGCTCTCTCAGCAGCAAGCAATTTTATGACCTGGTACAGTTCTTCTGAC	3854
Db	3481	AATTCAGATGCTCTCTCAGCAGCAAGCAATTTTATGACCTGGTACAGTTCTTCTGAC	3540
QY	3855	TTTCTTTCAGTTGTTCTCTCTTTTCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3914
Db	3541	TTTCTTTCAGTTGTTCTCTCTTTTCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3600
QY	3915	AAGTTGTTGGGCTCCAGAGCTAGAGAGCAACCTTTTCAGAGATGCAACAGATGGCCAG	3974
Db	3601	AAGTTGTTGGGCTCCAGAGCTAGAGAGCAACCTTTTCAGAGATGCAACAGATGGCCAG	3660
QY	3975	CGTCCCTTTGCTGTAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4034
Db	3661	CGTCCCTTTGCTGTAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3720
QY	4035	TGGGGGAGTATAAAGACTGTTCCCAACCCATTGCACTGGAGCCGCTGTTTGGCAACAA	4094
Db	3721	TGGGGGAGTATAAAGACTGTTCCCAACCCATTGCACTGGAGCCGCTGTTTGGCAACAA	3780
QY	4095	AGCCGCTGCTCTCTGCTGTTGAGGCTCCCTCGAGGCTGGGCTGCTGCTGCTGCTGCTGCT	4154
Db	3781	AGCCGCTGCTCTCTGCTGTTGAGGCTCCCTCGAGGCTGGGCTGCTGCTGCTGCTGCTGCT	3840
QY	4155	TGGGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4214
Db	3841	TGGGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3900
QY	4215	AGTGTACAGGAGAGTCAAGAGCCGCTGAGAAACCGGAGAGCAGAGCCCTGTCAGAGCC	4274
Db	3901	AGTGTACAGGAGAGTCAAGAGCCGCTGAGAAACCGGAGAGCAGAGCCCTGTCAGAGCC	3960
QY	4275	TGGGACCTGCTGTA 4290	
Db	3961	TGGGACCTGCTGTA 3976	

RESULT 2
AY418589
LOCUS
DEFINITION Mus musculus ATN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

3976 bp DNA linear GSS 17-DEC-2003

ACCESSION AY418589
 VERSION GI:39774549
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 3976)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D., and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBLISHED 14671302
 REFERENCE 2 (bases 1 to 3976)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D., and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 source location/Qualifiers
 1..3976
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1..>3976
 /gene="ATRN"
 /locus_tag="HCM6607"
 ORIGIN
 Query Match 68.1%; Score 2923.6; DB 29; Length 3976;
 Best Local Similarity. 78.3%; Pred. No. 0;
 Matches 3115; Conservative 0; Mismatches 861; Indels 0; Gaps 0;
 QY 315 TGTCAACGGCGTCTGCTGCAACCTCGGACCGGCGAGTGCCTCTGCCCCCGCGCTGGT 374
 DB 1 TGTCAACGGCGCGCTGCAACCTCGGACCGGCGAGTGCCTCTGCCCCCGCGCTGGT 60
 QY 375 GGGCAGCAATGCCAGCACTCGGGGGCGCTTCAGACTACTGATCTCTCGGGTTGT 434
 DB 61 GGGCAGCAATGCCAGCACTCGGGGGCGCTTCAGNNNNNNNNNNNNNNNNNNNNNN 120
 QY 435 GACAGATGGACCTGGAATTATAAATAACAAACGAAAGTGCACGTGGCTCATTTGAAGGACA 494
 DB 121 NNN 180
 QY 495 GCCAAATAGATATGAGACTTCGTTCAATCATTTTGTACAGAGTGTAGTTGGGACCA 554
 DB 181 GCCAAATAGATATGAGACTTCGTTCAACCAATTTTGTACAGAGTGTAGTTGGGACCA 240
 QY 555 TTTATATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTCTGCTCATTTAGTGGCCT 614
 DB 241 TTTATATGTTTATGATGGGACTCAATCTACGCACTCTGATTCGCTTTAGNNNNNN 300
 QY 615 CATTTCTCCTGAGAGATGCAATGAGACTGTCCCTGAGTGTGTTGCCACATCAGTTA 674
 DB 301 NNN 360
 QY 675 TGCCTTGTGCTATTTTATGATGCTGCTTATAAATTCAGTGGATTTAATTAATCTTA 734
 DB 361 NNN 420
 QY 735 CAGTTTGTATGTTGTCATAATACCTGTCAGGCCGAGAGAGTGTAAAGATCAGTAATAG 794
 DB 421 NNNTTTTGTATGTTGTCGGAATATTTGCTCAGGCCGAGAGAGTGTAAAGATCAGTAATAG 480
 QY 795 CAGCGAAACTGTTTCAATGTGAATGTTCTGTAAGAACTGGAAGAGTGAAGCATGTGACATTC 854

DB 481 CAGCAGCGCTGTTGAGTGTGAATGTTCTGAAACTGGAAGGGAGTCTGTGACATTC 540
 QY 855 TCACGTGACAGACAACCTGTGGTTTCTCATCAGGCACTGCAATTCAGTGTATGTCAG 914
 DB 541 TCACGTGACAGACAACCTGTGGCTTCTCACCAGGCACTGTGTAATGCAAGCATACCA 600
 QY 915 AGGATGCTCTGCTCTTCTCAGACTGGCAGGCTCCTGATGTTGATGTTCTGTACCACTAA 974
 DB 601 AGGATGCTCTGCTCTTCTCAGCTGGCAGGCTCCTGATGTTGATGTTCTGTGCGCACTAA 860
 QY 975 CAGTCAATTTGAGCTGCGAGAGGAAATATCTAACTTAAAGCTCCCGAGAGCACTCATAA 1034
 DB 661 CCACTCTTTTGGACTCGAGAGAAATATCTGATTTAAAGCTTCCAGAGCCTCTCATAA 720
 QY 1035 AGCTGTGCTCAATGGAACCAATATATGTTGGTGTGAGGATATATGTTCAACCACTCAGA 1094
 DB 721 AGCTGTGCTCAATGGAACCAATATATGTTGGTGTGAGGATATATGTTCAACCACTCAGA 780
 QY 1095 TTATAACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGTTTC 1154
 DB 781 TTACAGCATGGTTTATAGCGTATGACCTGACCTTCTAGGAAATGGCTTCCACTAAACCATTC 840
 QY 1155 TGTGAACAATGTTGTTAGATATGATATGTTCTTCTGTCATTTGGCAATATACAGAGATAAATTTA 1214
 DB 841 TGTGAACAATGTTGTTAGATATGATATGTTCTTCTGTCATTTTGGCAATATACAGAGATAAATTTA 900
 QY 1215 CATGTATGAGGAGAAATTTGATCTCAACTGGGAATGACCAATGAGTTCAGAGATTTTTC 1274
 DB 901 NNN 960
 QY 1275 CATTCAATGAGTCTATGGGTGTTGTGACCCCTAAGGCAAGAGAGTATGCAAGTGT 1334
 DB 961 NNN 1020
 QY 1335 TGGGCACTCTGCACACATTTGTACACTGAAGAATGGCCGAGTGTGTCATGCTGCTCATCTT 1394
 DB 1021 NNN 1080
 QY 1395 TGTCTACTGCCCTCTCTATGATATATAGCAATGTGCAGGAATATGATTTGGAATAGAA 1454
 DB 1081 NNN 1140
 QY 1455 CACATGAGTATATTACACACCCAGGCTGCTCTGCAAGGGGTTACGGCCATAGCAG 1514
 DB 1141 CACATGAGTATATTACATCTCAGGCTGCTCTGCAAGGGGTTATGGCCACAGTAG 1200
 QY 1515 TGTTCACGACCATAGGACCAAGGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1574
 DB 1201 TGTTCATGATGACAGGACCAAGGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 QY 1575 GAATTAAGTACCGGCTGACAGATCTCTACGATATATGATGATGATGATGATGATGATGATGAT 1634
 DB 1261 CAACAATACCGGCTGACAGATCTCTACGATATATGATGATGATGATGATGATGATGATGAT 1320
 QY 1635 CATTCTTAAGGACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1694
 DB 1321 CATTCTTAAGGACACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 1695 CATGCTGTTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGCGGCCAAATG 1754
 DB 1381 CATGCTGTTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGCGGCCAAATG 1440
 QY 1755 CTTCTCTTCAATTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814
 DB 1441 CTTCTCTTCAATTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1815 ACCTGATCTCCACCATGATGCAACAGATTTGGCCATTCAGCAGTCTTTACACACACAC 1874
 DB 1501 NNN 1560
 QY 1875 CATGTATGTTGTTGGGTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1934

1561 CATGTATGTGTTGGCGGCTTCAACAGCCCTCCTCCTCAGTGACGCTCTTGGTCTTTACCTC 1620
1935 GGAACAGTGTGATGGGATCGAGTGAAGCGCTGTTTGTAGCAGCAGGACTGGTATTCG 1994
1621 GGAGCGTGGCATGACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACTGGTATCCG 1680
1995 GTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACA 2054
1681 GTGTGTGTGGACACACAGAGTGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACA 1740
2055 AGAGAAAGTTAAATCAGAAATGTTTTCAAAGAACTCTTGACCATGACAGATGTA 2114
1741 AGCAGAAAGTTAAATCAGAGTGTGTTTCTTAAAGAAACCCCTTGACCATGACAGATGTA 1800
2115 CCAGCACACAGATGTTTACAGCTGTACAGGCCAACCAATCAATCACTGCCACTGGTCCAATGA 2174
1801 CACGACACAGATGTTTACAGCTGTACAGGCCAACCAATCAATCACTGCCACTGGTCCAATGA 1860
2175 CCAATGTCTCCCAAGGAACACAGCTGTCTAGAGGCCAGATCTCCATTTTATGGTATGA 2234
1861 TCACGTGTCTGTGAACACAGCTGTGACAGAGGCCAGATCTCCATTTGCCAAGTATGA 1920
2235 GAAATGCCCCAAGGATAACCCCATGTACTGTAAACAGAGAACCCAGCTGCAGAGCTG 2294
1921 GAGTTGCCCCCAAGGATAACCCCATGTACTGTGCNATGAAGAAACCCAGCTGCAGAGCTG 1980
2295 TGCCCTGGACACAGAACTGCCAGTGGAGCCCGGCAATCAGGAGTGCAATGGCCCTGCCCGA 2354
1981 TGCCCTAGACAGAACTGCCAGTGGAGCCCGGCAATCAAGAGTGCAATGCCCTGCCCGA 2040
2355 AAATATCTGTGGCATGTGCTGGCATTTGGTTGGAACCTCATGTTTGAATAATCTACTGC 2414
2041 AAATATCTGTGGCAATGTGCTGGCATTTGGTTGGAACCTCGTGTCTGAATAATCACTACTGC 2100
2415 CAAGAGAGAAATATGACAAATGCTTAAATGCTCTGTAGGAACACCAATGCCCCCTTTTGGCTTC 2474
2101 TAAAGAGAAATATGACAAATGCTTAAATGCTCTGTAGGAACACCAATGCCCCCTTTTGGCTTC 2160
2475 TCTTAAACCCCAAGAGAGTAGAATTTGCTCTTAAGCAGCTGGCAATGAGTGCATC 2534
2161 CCTCACATCCCAAGAGAGTAGAATTTGCTCTTAAGCAGCTTCGATTAATGCAATCATC 2220
2535 TCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGCTCTCGAAGAGTCAATGTGTC 2594
2221 TCAAGATATGCTCAAGCTCACTCTCACTCCATGGGTGGTCTTTCGGAAGATCAATGTGTC 2280
2595 CTACTGGTCTGGGAAGATATGCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2654
2281 TTACTGGTCTGGGAAGATATGCTCCCATTTACAAATAGTTTGTCTGACAGTGGATGCCATC 2340
2655 TGAGCCCATGATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2714
2341 TGAGCCCATGATGCTGGATTTCTGTGGATCTTGTGAGGCTAGTACTCGGGGAAATAAA 2400
2715 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCCTGCAAAACCAAG 2774
2401 GGCTGCAACCTGCATCAACCCCTCTCAATGGCAGCTCTGTGAAGGCCCTGCAAAACCAAG 2460
2775 TGCTAAGCAGTCCCGACACCAATGTGCTTGAAGACAGCATGTGAGATTTGACAGCGG 2834
2461 TGCCAGCAGTCCCGACACCAATGTGCTTGGGACAGCGTGTGGCGAGTGCACCTAGCAG 2520
2835 CAGCTCTGATGCTGTGGTGCAGCAATGAAGCAGTGTGTGGACTCCAATGCCATGT 2894
2521 CAGCTCTGAGTGCATGTGGTGCAGTAAATGAAGCAGTGTGTGGACTCCAATGCCATGT 2580
2895 GGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGACCACTGCCCCCTTGA 2954
2581 GGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGACCACTGCCCCCTTGA 2640
2955 AAATTTCTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGAGCTGTGGCTGGT 3014
2641 AAATTTCTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGAGCTGTGGTGGT 2700

3015 TACTGATCCAGCAATACTGGCAAAAGGAAATGCATPAGAGGTTCTTATAAAGGACCACT 3074
2701 TACTGATCTAGCAATACTGGCAAAAGGAAATGTAATGAGGCGAGCTATAAAGGACCTGT 2760
3075 GAAGATGCTTGCAGAGCCCTTACAGGAAATTTCTATCCAGAGCCCTCTCAATTCAG 3134
2761 GAAGATGCTTGCAGAGCCCTTGCAGGAAATGTATCCAGAGCCCTTCTGAACCTCCAG 2820
3135 CATGTGCTTAGAGGACAGCAGATACAACTGGTCTTTTCACTGTCTCCAGCTTGCCAAATG 3194
2821 CATGTGCTTAGAGGACAGCAGATACAACTGGTCTTTTCACTGTCTCCAGCTTGCCAGTG 2880
3195 CAAAGGCCACAGTAATGCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTTGACCAAC 3254
2881 CAAAGGCCACAGTAATGCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTTGACCAAC 2940
3255 AGCAGAGCACTGGAGACCTGCATATCTGCTTCTACGCTGTATCCACCAATGGAGGAA 3314
2941 GGCAAGCACTGGAGACCTGCATATCTGCTTCTATGCTGACCCGACTATATGGAGGCAA 3000
3315 ATGTAGCCATGCAAGTGCATGGGCACTGTCTGTGCAACCAACACAGCGGCAAGTG 3374
3001 ATGTAGCCATGCAAGTGCATGGGCACTGTCTGCAACCAACCAACAGCGGCAAGTG 3060
3375 CTTTGTGACCAACAGGGGCTCAAGGGGACAGTGCAGCTATGTAGGTAGAAATCG 3434
3061 CTTTGTGACCAACAGGGGCTCAAGGGGACAGTGCAGCTATGTAGGTAGAAATCG 3120
3435 ATACCAAGGAAACCCCTCTCAGAGGAACTGTATTATCTCTTCTTATTGACTATCAGTT 3494
3121 ATACCAAGGAAACCCCTCTCAGAGGAACTGTACTATACCTTCTCATTGACTATCAGTT 3180
3495 CACTTTAGTCTATCCAGAGGATGATCGCTATTACAGAGCTATCAATTTTGTGCTAC 3554
3181 CACTTTAGCTCTCCAGAGGACAGCTACTACTACAGAGCTATCAACTTTGTGCTAC 3240
3555 TCCTGAGCAACAAACAGGGATTTGACATGTTCACTCAATGCTCCAAAGAACTTCAACT 3614
3241 TCCTGATGAACAAACAGGGATTTGACATGTTCACTCAATGCTCCAAAGAACTTCAACT 3300
3615 CAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGT 3674
3301 CAACATCAGCTGGGCTGCCAGCTTCCAGCGGAAACCCAGACTGGAGAGAGGCTGTGT 3360
3675 TGTTTCAAAACCAACATTAAGAGTACAAAGATGTTTCTCTAATGAGAGATTTGATTT 3734
3361 TGTTTCAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTT 3420
3735 TCGCAACCCCAATATCACTTTTCTTTTGTATGTCACTAATTTTCACTGGCCCATCAA 3794
3421 TCGCAACCCATCCAAACATCACTTTTCTTTTGTATGTCACTAATTTTCACTGGCCCATCAA 3480
3795 AATTGAGATTCCTTCTCAGCAGCAATTTTATGACCTGCTGACAGTCTTCTGCTGAC 3854
3481 AATTGAGATTCCTTCTCAGCAGCACTTCCAGCAGCACTTCACTGGACCTGCTGCTGAC 3540
3855 TTTTCTCAGTTGTTTCTCTCTCTTTGCTCTGCTGGTCTGCTGTTTGAAGATCAAACA 3914
3541 TTTTCTCAGTTGTTTCTCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
3915 AAGTTGTGGGCTTCCAGAGCTAGAGAGCAATTTCTTGGAGATGCAACAGATGGCCAG 3974
3601 GAGCTGTGGGCTTCCAGGCGGAGAGAGCACTTCTTGGGAGATGCAACAGATGGCCAG 3660
3975 CCGTCCCTTTGCTCTGTAATGCTGCTTGGAAACAGATGAGAGCTCTCTGATCTTAT 4034
3661 CCGGCTCTTGTCTGTAACAGTGTGCTTGGAAACAGATGAGAGCTCTCTGATCTCAT 3720
4035 TGGGGGAGTATAAAGACTGTTCCTCAACCCATTTGCACTGGAGCCGTTGTTTGGCAACA 4094
3721 TGGGGGAGTATAAAGACTGTTCCTAAGCCCAATGCTTGGAGCCCTGCTTGGTAACA 3780

1755 QY CTTCTCTTCCAGATTTCATGSCCTTATGACATTTGCTGTGACCGCTGCTCAGTGTCTCCAG 1814
1261 Db CTTCTCTTCCAGATTTCATGSCCTTATGACATTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1320
1815 QY ACTGATCTCCACCATGATGTCAACAGATTGGCCATTGAGCAGTCTTACACACAGCAG 1874
1321 Db NNN 1380
1875 QY CATGTATGTCTTGGTGGTTCATATAGTCTCTCCCTCAGCGACATCTCGTATTCACCTC 1934
1381 Db CATGTATGTCTTGGTGGTTCATATAGTCTCTCCCTCAGCGACATCTCGTATTCACCTC 1440
1935 QY GGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTATTAGCAGCAGGACCTGGTATTCG 1994
1441 Db GGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTATTAGCAGCAGGACCTGGTATTCG 1500
1995 QY GTGTGTGTGAACACAGCGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTCATGAACA 2054
1501 Db GTGTGTGTGAACACAGCGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTCATGAACA 1560
2055 QY AGAAGAAAGTTAAATCAGAAATGTTTTTCAAAGAACTCTTGACCATGACAGATGTGA 2114
1561 Db AGAAGAAAGTTAAATCAGAAATGTTTTTCAAAGAACTCTTGACCATGACAGATGTGA 1620
2115 QY CCAGCACAGATGTTACAGCTGTACAGCCACACCAATGACTGCCACTGGTGCATGA 2174
1621 Db CCAGCACAGATGTTACAGCTGTACAGCCACACCAATGACTGCCACTGGTGCATGA 1680
2175 QY CCATTGTGTCTCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTGGTATGA 2234
1681 Db CCATTGTGTCTCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTGGTATGA 1740
2235 QY GAATTTGCCCCAGGATTAACCCCATGTACTGTCTAAACAGAGACAGCTGCGAGGAGCTG 2294
1741 Db GAATTTGCCCCAGGATTAACCCCATGTACTGTCTAAACAGAGACAGCTGCGAGGAGCTG 1800
2295 QY TGCCCTGACACAGAACTGCCAGTGGGAGCCCGGAACTCAGAGTGCATTTGCCCTGCCCGA 2354
1801 Db TGCNNNGACACAGAACTGCCAGTGGGAGCCCGGAACTCAGAGTGCATTTGCCCTGCCCGA 1860
2355 QY AAATATCTGTGGCAATTTGCTGGCAATTTGTTGGAAATCTCATGTTTGAATTTACTATGTC 2414
1861 Db AAATATCTGTGGCAATTTGCTGGCAATTTGTTGGAAATCTCATGTTTGAATTTACTATGTC 1920
2415 QY CAAGGAGAAATATGCAATGTCTAAATTTCTGTAGGAACACCAATGCCCTTTTGGGTTTC 2474
1921 Db CAAGGAGAAATATGCAATGTCTAAATTTCTGTAGGAACACCAATGCCCTTTTGGGTTTC 1980
2475 QY TCTTACAAACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTGCGAAATTAATGCAATCATC 2534
1981 Db TCTTACAAACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTGCGAAATTAATGCAATCATC 2040
2535 QY TCAGAGCATGTCAGAGTCACTTAACCCCATGGTGGGCTTGGAGATCAATGTGTC 2594
2041 Db TCNNNNNNNTCCAGAGCTCACCTTAACCCCATGGTGGGCTTGGAGATCAATGTGTC 2100
2595 QY CTACTGGTGTGGGAAGATATGCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2654
2101 Db CTACTGGTGTGGGAAGATATGCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2160
2655 QY TCAGCCCAAGTGTGATGTTGGAATTTTATCAGAAACCCAGTACTCGGAGACTGAA 2714
2161 Db TCAGCCCAAGTGTGATGTTGGAATTTTATCAGAAACCCAGTACTCGGAGACTGAA 2220
2715 QY GGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGAAAGCCCTGCAAAACACAG 2774
2221 Db GGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGAAAGCCCTGCAAAACACAG 2280
2775 QY TGCTAAGCAGTGCAGACACCAATGTGCTTGGAGCAGCATGTGGAGATTCACAGCGG 2834
2281 Db TGCTAAGCAGTGCAGACACCAATGTGCTTGGAGCAGCATGTGGAGATTCACAGCGG 2340
2835 QY CAGCTCTGATGCTGTGTGAGCAACATGAGCAGTGTGTGACTCCATGCTCATGT 2894

2341 Db CAGCTCTGAGTGCAATGTGTGCGACCAATGAGCAGTGTGTGACTCCATGCTACGT 2400
2895 QY GGCCTCTCTCCCTTTTGGCCAGTGTATGGAATGTGTATACGATGAGCAGCTCCGCCCTCGA 2954
2401 Db GGCCTCTCTCCCTTTTGGCCAGTGTATGGAATGTGTATACGATGAGCAGCTCCGCCCTCGA 2460
2955 QY AAATTTGTTCAGGCTACTGTACTCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTG 3014
2461 Db NNN 2520
3015 QY TACTGATCCAGCAATACTGGCAAAGGGAATGTCATAGAGGGTTCTATAAAGGACCAGT 3074
2521 Db NNN 2580
3075 QY GAAGATGCTTCGCAAGCCCTACAGGAAAATTTCTATCCACAGCCCTGCTCAATTCACAG 3134
2581 Db NNN 2640
3135 QY CATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTCAATTCAGTGTCCAGCTTGCCTCAATG 3194
2641 Db NNN 2700
3195 QY CAAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCAAC 3254
2701 Db CAAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCAAC 2760
3255 QY AGGCAAGCACTCGGAGACCTGCATCTCTGGCTTCTACGGTGTATCCCACTCAATGGAGGAA 3314
2761 Db AGGCAAGCACTCGGAGACCTGCATCTCTGGCTTCTATGGTGTATCCCACTCAATGGAGGAA 2820
3315 QY ATGTACAGCCATCAAGTGCATTTGGGACGCTCTGTGCAACACCAACACGGGCAAGTG 3374
2821 Db NNN 2880
3375 QY CTTCTGCACACCAAGGGCTCAAGGGGACAGTGCAGCTATGTGAGTGAAGAAATCG 3434
2881 Db CTTCTGCACACCAAGGGCTCAANN 2940
3435 QY ATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTAATCTCTTCTTATTAATGACTATCAGTT 3494
2941 Db ATACNN 3000
3495 QY CACCTTTAGTCTATCCAGGAAGATGATCGTATTTACAGAGTATCAATTTTGTGGCTAC 3554
3001 Db CACCTTTAGTCTATCCAGGAAGATGATCGTATTTACAGAGTATTAATTTTGTGGCTAC 3060
3555 QY TCCTGACGACAAACACAGGATTTGGACATGTTCAATGCTCCAGAAATTTCAACCT 3614
3061 Db TCCTGACGACAAACACAGGATTTGGACATGTTCAATGCTCCAGAAATTTCAACCT 3120
3615 QY CAAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGT 3674
3121 Db CAAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGT 3180
3675 QY TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTT 3734
3181 Db TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTT 3240
3735 QY TCGCAACCCCAAAATACACTTTTCTTTGTTATGTAGTAAATTTTCACTGGGCCATCAA 3300
3241 Db TCGCAACCCCAAAATACACTTTTCTTTGTTATGTAGTAAATTTTCACTGGGCCATCAA 3360
3301 QY AATTGAGATGCTCTCTCAGCAGCAGCAATTTTATGGACCTGGTACAGTTCTTCGTGAC 3354
3355 QY AATTGAGATGCTCTCTCAGCAGCAGCAATTTTATGGACCTGGTACAGTTCTTCGTGAC 3360
3361 Db TTTCTTCAGTTGTTTCTCTCTTTTGTCTCTGCTGGCTGTGTTGGAAGATCAAAACA 3914
3915 QY AAGTTGTGGGCTTCCAGAGTGAAGACAACTTCTTCGAGAGATGCAACAGATGGCCAG 3974

```

Db      3421 AAGTTGTTGGGCTCCAGACGTAGAGCAACTCTTCGAGAGATGCAACAGATGGCCAG 3480
QY      CGTCCCTTTGGCTCTGTAATGTGGCTTGGCTTGGAAACAGATGAGGAGCTCTCTGATCTTAT 4034
Db      3481 CGTCCCTTTGGCTCTGTAATGTGGCTTGGCTTGGAAACAGATGAGGAGCTCTCTGATCTTAT 3540
QY      TCGGGGGAGTAAAGACTGTTCCCAACCCCAATTCACCTGAGCGGTGTTTGGCAACA 4094
Db      3541 TCGGGGGAGTAAAGACTGTTCCCAACCCCAATTCACCTGAGCGGTGTTTGGCAACA 3600
QY      ACCGCTGTCCTCTCTGTTGTCAGGCTCCCTCGAGGCTGAGGCTGAGGCTCCCTCTCC 4154
Db      3601 AGCCGCTGTCCTCTCTGTTGTCAGGCTCCCTCGAGGCTGAGGCTGAGGCTCCCTCTCC 3660
QY      4155 TGGGCGAGTCAG 4165
Db      3661 TGGGCGAGTCAG 3671

RESULT 4
LOCUS   BX440935
DEFINITION BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION BX440935
VERSION   BX440935.1 GI:30781889
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10212.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF012CF10QPI&cluster=10212.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF012CF10QPI.
FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF012YI19"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT.6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

Query Match 21.68; Score 925.6; DB 13; Length 1201;
Best Local Similarity 98.4%; Pred. No. 5.5e-166;
Matches 962; Conservative 4; Mismatches 8; Indels 4; Gaps 3;
QY 2337 GTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCATTTGGTGGAACTCATG 2396
Db 57 GTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCATTTGGTGGAACTCATG 116
QY 2397 TTTGAAATTAATCTGCCAAGGAGATTAATGCAATGTAATTTCTGTAGGAACA 2456

```

```

Db      117 TTTGAAATTAATCTGCCAAGGAGATTAATGCAATGTAATTTCTGTAGGAACA 176
QY      CAATGCCCCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGTCTTAAAGCAGCT 2516
Db      177 CAATGCCCCCTTTTGGCTTCTCTTACAAACCCCTGATGAGGTAGAAATTTGTCTTAAAGCAGCT 236
QY      GCAATTAATGCAATCTCTCAGAGCATGTCAGAGCTCACTTAACCCCATGAGGTGGGCT 2576
Db      237 GCAATTAATGCAATCTCTCAGAGCATGTCAGAGCTCACTTAACCCCATGAGGTGGGCT 296
QY      TCGGAAGATCAATGTCCTCTCTGTTGGTGGGAAGATATGCCCATTTTACAAATAGTTT 2636
Db      297 TCGGAAGATCAATGTCCTCTCTGTTGGTGGGAAGATATGCCCATTTTACAAATAGTTT 356
QY      ACTACAGTGGATGCCGTCTGAGCCCCAGTGAATCTGTGGAATTTTATCAGAACC 2696
Db      357 ACTACAGTGGATGCCGTCTGAGCCCCAGTGAATCTGTGGAATTTTATCAGAACC 416
QY      CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGA 2756
Db      417 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGA 476
QY      AAGGCTTGCAAAACCCAGTGTAAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCATG 2816
Db      477 AAGGCTTGCAAAACCCAGTGTAAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCATG 536
QY      TGGAGATTGCACAGGGGAGCTCTGAGTGCATGTGTCAGCAACATGAAGCAGTGT 2876
Db      537 TGGAGATTGCACAGGGGAGCTCTGAGTGCATGTGTCAGCAACATGAAGCAGTGT 596
QY      GGACTCCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACCAT 2936
Db      597 GGACTCCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACCAT 656
QY      GAGCACCTGCCCCCTTGAAATTTGTTAGGCTACTGTACTCTGATCTTCTTGGAGCA 2996
Db      657 GAGCACCTGCCCCCTTGAAATTTGTTAGGCTACTGTACTCTGATCTTCTTGGAGCA 716
QY      ACCAGCTGTGGCTGTGCTGCTACTGATCCAGCAATCTGCGAAAGGAATCCATAGAGGG 3056
Db      717 ACCAGCTGTGGCTGTGCTGCTACTGATCCAGCAATCTGCGAAAGGAATCCATAGAGGG 776
QY      TTCTTATAAAGGACCAAGTGAAGTCCCTTCGCAAGCCCTTACAGGAAATTTCTATCCACA 3116
Db      777 TTCTTATAAAGGACCAAGTGAAGTCCCTTCGCAAGCCCTTACAGGAAATTTCTATCCACA 836
QY      GCGCTGCTCAATTCACAGTGTCTGTAGAGGACAGAGATACACTGGCTTTTCAATCA 3176
Db      837 GCGCTGCTCAATTCACAGTGTCTGTAGAGGACAGAGATACACTGGCTTTTCAATCA 896
QY      CTGTCCAGCTTGCCTAATGCAACGGGCCACAGTAATGCAATCAGAGCATCTCTGAGAA 3236
Db      897 CTGTCCAGCTTGCCTAATGCAACGGGCCACAGTAATGCAATCAGAGCATCTCTGAGAA 956
QY      GTGTGAGAACCTGACCACAGGCAAGCTGCGAGCTGCGATATCTGGCTTCTTACGGTGA 3296
Db      957 GTGTGAGAACCTGACCACAGGCAAGCTGCGAGCTGCGATATCTGGCTTCTTACGGTGA 1013
QY      TCCCAACCAATGGAGGAA 3314
Db      1014 T-CCACCAATGGAGGAA 1030

RESULT 5
LOCUS   BU506373
DEFINITION AGENCOURT 10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:649441
ACCESSION BU506373
VERSION   BU506373.1 GI:22812606
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 964)

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM14050 row: k column: 18
High quality sequence stop: 592.
Location/Qualifiers
1. 964
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:649441"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 18.1%; Score 775; DB 13; Length 964;
Best Local Similarity 88.9%; Pred. No. 3e-137;
Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;

QY 1996 TGTGTGGAAACACAGGTCCTCTCAGTGTATCTGTCGGGGCTGGCAACTGTATGACAA 2055
DB 1 TGTGTGGGACACACAGTCGTCTCGATGTACCTCTCTGGAGTTGGCAACTGGAACAA 60

QY 2056 GAAGAAAGTTAAATCAGATGTTTCCAAAGACTCTTGACCATGACAGATGAC 2115
DB 61 GCAGAAAGTTAAATCAGATGTTTCTTAAAGAACCTTGACCATGACAGATGAC 120

QY 2116 CAGCACAGATTTGTACAGCTGTACAGCCACACCAATGATCCACTGGTGGAAATGAC 2175
DB 121 CAGCACAGATTTGTACAGCTGTACAGCCACCAATGATCCACTGGTGGAAATGAC 180

QY 2176 CATTGTGCCCGAAGAACACACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTATGAC 2235
DB 181 CACTGTGTCCTGTGAACACACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTATGAC 240

QY 2236 AATTGCCCGAAGATAACCCATGTACTGTCTTAAAGAACCTTGACCATGACAGATGAC 2295
DB 241 AGTTGCCCGAAGATAACCCATGTACTGTCTTAAAGAACCTTGACCATGACAGATGAC 300

QY 2296 GCGTGGACCAAGATGCGAGTGGAGGCCCGGAATCAGAGTGAATGCGCTGCCGAA 2355
DB 301 GCGTGGACCAAGATGCGAGTGGAGGCCCGGAATCAGAGTGAATGCGCTGCCGAA 360

QY 2356 AATATCTGGCATTTGGTGGAACTATGTTGAAACTATGTTGAAACTATGTTGAACTATGTTG 2415
DB 361 AATATCTGGCATTTGGTGGAACTATGTTGAAACTATGTTGAAACTATGTTGAACTATGTTG 420

QY 2416 AAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAATGCGCTTTGGCTTCT 2475
DB 421 AAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAATGCGCTTTGGCTTCT 480

QY 2476 CTTACAAACCAAGAGTGAATTTGCTTACGAGCTCGGAATGAGTCACT 2535
DB 481 CTCACATCCAGAGAGTGGAGTTGTCTTAAAGAGCTTCGATTATGCAATCATCT 540

2536 CAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGCGCCTTCGGAAGATCAATGTGTCC 2595
DB 541 CAAAGTATGTCCAAGCTCACCTGTGACTCCATGGGTGGTCTTCGGAAGATCAATGTGTCT 600

QY 2596 TACTGTGCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCCGTCT 2655
DB 601 TACTGTGCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTCTGCAATGATGCCATCT 660

QY 2656 GAGCCCAAGTGTCTGGGATTTCTGTGGAATTTTATCAGAACCCAGTACTC-GGGAGCTGAA 2714
DB 661 GAGCCCAAGTGTCTGGGATTTCTGTGGAATTTTATCAGAACCCAGTACTC-GGGAGCTGAA 720

QY 2715 GGCTCAACCTGTCATCAACCACTCAATGATGTCTGTGAAAGGCTGCAAAACACAG 2774
DB 721 GGCTCAACCTGTCATCAACCACTCAATGATGTCTGTGAAAGGCTGCAAAACACAG 780

QY 2775 TGCTAAGCAGTGGCGGACACCATGTGCTTGAGGACAG-CATGTGGAGATTGCACACGC 2833
DB 781 TGCTAAGCAGTGGCGGACACCATGTGCTTGAGGACAG-CATGTGGAGATTGCACACGC 840

QY 2834 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGTGACTTCCATGCTATG 2893
DB 841 CCAGCTCGGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGTGACTTCCATGCTATG 900

QY 2894 TGGCT 2953
DB 901 TGGCT 960

QY 2954 AAAA 2957
DB 961 AAAA 964

RESULT 6
EG678679 933 bp mRNA linear EST 01-MAY-2001
LOCUS 602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
DEFINITION mRNA sequence.
ACCESSION EG678679.1 GI:13910076
VERSION EG678679.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM10602 row: j column: 20
High quality sequence stop: 762.
Location/Qualifiers
1. 933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: oligo dr. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

ORIGIN

Query Match 17.7%; Score 758.8; DB 12; Length 933;
 Best Local Similarity 98.4%; Pred. No. 3.6e-134;
 Matches 777; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 2997 ACAGGCTGTGGCTGTGTACTGATCCAGCAATATGCGCAAGGGAATGATAGAGGG 3056
 DB 2 ACAGGCTGTGGCTGTGTACTGATCCAGCAATATGCGCAAGGGAATGATAGAGGG 61

QY 3057 TTCCTATAAGGACGAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 3116
 DB 62 TTCCTATAAGGACGAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 121

QY 3117 GCGCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCATTC 3176
 DB 122 GCGCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCATTC 181

QY 3177 CTGTCCAGTTCGCAATGCAAGCGCCACAGTAATGCAATCAATCAGACATCTGTGAA 3236
 DB 182 CTGTCCAGTTCGCAATGCAAGCGCCACAGTAATGCAATCAATCAGACATCTGTGAA 241

QY 3237 GTGTGAGAACCTTGACACAGGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGA 3296
 DB 242 GTGTGAGAACCTTGACACAGGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGA 301

QY 3297 TCCCAACCAATGAGGGAATGTGAGCAATGCAAGTGCATGCGCAAGCGCTCTGTGCAA 3356
 DB 302 TCCCAACCAATGAGGGAATGTGAGCAATGCAAGTGCATGCGCAAGCGCTCTGTGCAA 361

QY 3357 CACCAACAGGCAAGTGTCTGACCAACCAAGGCGTCAAGGGGACGATCGCACT 3416
 DB 362 CACCAACAGGCAAGTGTCTGACCAACCAAGGCGTCAAGGGGACGATCGCACT 421

QY 3417 ATGTGAGTGAAGAAATGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTACTCT 3476
 DB 422 ATGTGAGTGAAGAAATGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTACTCT 481

QY 3477 TCTTATTCAGTATCAGTTCAGCTTTAGTCTATCCAGGAAGATGCTATTACAGC 3536
 DB 482 TCTTATTCAGTATCAGTTCAGCTTTAGTCTATCCAGGAAGATGCTATTACAGC 541

QY 3537 TATCAATTTTGTGGCTACTCTGACGAAACCAACAGGATTTGACATGTTTCATCATGC 3596
 DB 542 TATCAATTTTGTGGCTACTCTGACGAAACCAACAGGATTTGACATGTTTCATCATGC 601

QY 3597 CTCGAAGATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTTCAGCTGGAACCCAGC 3656
 DB 602 CTCGAAGATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTTCAGCTGGAACCCAGC 661

QY 3657 TGGAGAGAGATGCTGTGTTTTCAGAAACCAACATTAAGAGTACAAAGATGTTCTC 3716
 DB 662 TGGAGAGAGATGCTGTGTTTTCAGAAACCAACATTAAGAGTACAAAGATGTTCTC 721

QY 3717 TAATGAGAAGTTGATTTTGCACACCAACCAATATCATCTTTCTTTGTTTATGTCAGTAA 3776
 DB 722 TAATGAGAAGTTGATTTTGCACACCAACCAATATCATCTTTCTTTGTTTATGTCAGTAA 780

QY 3777 TTTTCACTGG 3786
 DB 781 TTGTCACTTG 790

RESULT 7
 CB520374
 LOCUS
 DEFINITION UT-M-G10-cei-1-18-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
 IMAGE:6940067 5', mRNA sequence.
 ACCESSION CB520374.1 GI:29353729
 VERSION CB520374
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 840)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES
 source
 1..840
 Location/Qualifiers

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6940067"
 /issue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP G10"
 /note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according
 Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is ACCGAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 16.3%; Score 701.2; DB 14; Length 840;
 Best Local Similarity 90.2%; Pred. No. 3.5e-123;
 Matches 759; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 2151 CAATGACTGCCACTGGTCAATGACCATGTGTCCCAAGGATACCCCATGTACTACTGTA 2210
 DB 1 CAATGACTGCCACTGGTCAATGACCATGTGTCCCTGTGAACCAAGCTGCAGAGG 60

QY 2211 CCAGATCTCCATTTTAGGTATGAAATGCCCCCAAGGATACCCCATGTACTACTGTA 2270
 DB 61 CCAGATCTCCATTTAGGTATGAAATGCCCCCAAGGATACCCCATGTACTACTGTA 120

QY 2271 CAGAAGACAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGCCCCGAA 2330
 DB 121 TAAGAAAACCAAGCTGCAGAGCTGTGCCCTAGACAGAACTGCCAGTGGGAGCCCCGAA 180

QY 2331 TCAGAGTGCATTCGCCCTGCCGAAATATCTGTGGCAATGCTGGCAATGCTGGTGA 2390
 DB 181 TCAGAGTGCATTCGCCCTGCCGAAATATCTGTGGCAATGCTGGCAATGCTGGTGA 240

QY 2391 CTCATGTTGAAATTTACTACTGCAAGGAAATATGCAATGCTAAATTTCTCTAG 2450
 DB 241 CTCGTTGCTGAAATTTACTACTGCAAGGAAATATGCAATGCTAAATTTCTCTAG 300

QY 2451 GAACCAATGCCCTTTTGGCTTCTCTACAAACCAAGAGTGAATTTGTCCTTAA 2510
 DB 301 GAACCAATGCCCTTTTGGCTTCTCTACAAACCAAGAGTGAATTTGTCCTTAA 360

2511	GCAGCTCGCAATTAATCAGTGCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGT	2570
361	CGAGCTTCGATTAAATGCAATCATCTCAAAGTATGTCCAAAGCTCACTCTGATCCATCGGT	420
2571	CGGCTTCGGAAGATCAATGTGTCTCTACTGTGTCTGGGAAGATATGTCCCCATTTACAAA	2630
421	TGGTCTTCGGAAGATCAATGTGTCTTACTTGGTCTCTGGAGGATATGTCTCCATTCACAAA	480
2631	TAGTTTACTACGTGGATCGGCTCTGAGCCCATGATCTCGAATCTGTGTGAATTTTATC	2690
481	TAGTTTCTGCTGAGTGGATGCCATCTGAGCCCATGATCTGCTGTCTGTGGGATCTTGTC	540
2691	AGRACCCAGTACTCGGGAGCTGAAGGCTGCACCTTGCATCAACCCACCTCAATGSGTAGTGT	2750
541	AGACCTTAGTACTCTGGGGATTAAAGGCTGCACTTGCATCAACCTCTCAATGSCACGGT	600
2751	CTGTGAAAGCCTGC AAAACA CAGTCTTAAGCAGTGC CGGACA CCAATGTGCTTGAGGAC	2810
601	CTGTGAAAGCCTGC AANACCA CAGTGC CCAAGCAGTGC CGGACA CCAATGTGCTTGCGGAC	660
2811	AGCATGTGGAGATTGCACAGCGGAGCTCTCAGTGTGCATGTGGTCAGCAACAACATGAAGCA	2870
661	AGCGTGTGGCGAGTGC ACTAGCAGCAGCTCGGATGTCATGTGGTGCNAGTACATGAAGCA	720
2871	CTGTGTGGACTCCAATGCCATATGTGGCTTCCTTTCCCTTTTGGCCAGTGTATGGATGGTA	2930
721	TGTGTGTGGACTCCAATGCCATACGTGGCTTCCTTTCCCTTTTGGCCAGTGTATGGATGGTA	780
2931	TACCATGAGCACCTTGCCCCCTGAAAATTTGTTCCAGGCTACTGTACTCTGTAGTCATTGCTT	2990
781	TACATGAGCAGCTGCCCACTTG - ANATTGCTCTGCGTACTGTACTCTGAGCAATTGCTT	839
2991	G 2991	
840	G 840	
Db		

RESULT 8	BM783739	LOCUS	681 bp	mRNA	linear	EST 05-MAR-2002
		DEFINITION	BM783739	S55NU484	Homo sapiens	CDNA clone S55NU484-34-H05 5', mRNA sequence.

BM783739
 BM783739.1 GI:19131971
 EST.
 Homo sapiens (human)
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 681)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE
21C Frontier Korean EST Project 2001
JOURNAL
Unpublished (2002)
COMMENT
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel : +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: H column: 05
High quality sequence stop: 681.

FEATURES	source
Location/Qualifiers	
1. .681	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="S55NU484-34-H05"	
/sex="M"	
/tissue_type="Stomach"	
/cell_type="Epithelial"	

RESULT 9
BM783788

/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/site="Organ: Stomach; Vector: pTZ19p1; Site1: EcoRI;
Site2: NotI; (A) + RNA was deccapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-tailed
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
method." method="electroporation
method."

ORIGIN

Query Match	15.9%;	Score 681;	DB 12;	Length 681;
Best Local Similarity	100.0%;	Prod. No. 2.3e-119;	Mismatches 0;	Gaps 0;
Matches 681;	Conservative			
2512	CAGCTGCGAATAATGCAGTCAATCTCAGAGATGTCCAAAGCTCACCTTAAACCCCATGGGTC	2571		
1	CAGCTGCGAATAATGCAGTCAATCTCAGAGATGTCCAAAGCTCACCTTAAACCCCATGGGTC	60		
2572	GGCCTTTGGGAAGATCAATGTGTCTACTGTGTCGTGGGAAGATATGTCCCATTTACAAT	2631		
61	GGCCTTTGGGAAGATCAATGTGTCTACTGTGTCGTGGGAAGATATGTCCCATTTACAAT	120		
2632	AGTTTACTACAGTGAGTGCCTCTGAGCCGACATGCTGGATCTGTGGAAATTTATCA	2691		
121	AGTTTACTACAGTGAGTGCCTCTGAGCCGACATGCTGGATCTGTGGAAATTTATCA	180		
2692	GAACCCAGTACTCGGGGACTGAAGCTGCAACCTTGCACTCAACCCACTCAATGGTAGTGTCT	2751		
181	GAACCCAGTACTCGGGGACTGAAGCTGCAACCTTGCACTCAACCCACTCAATGGTAGTGTCT	240		
2752	TGTGAAAGCCCTGCAAAACCAACAGTCTTAAGCAGTCCCGGACACCATGTGCTTTGAGGACA	2811		
241	TGTGAAAGCCCTGCAAAACCAACAGTCTTAAGCAGTCCCGGACACCATGTGCTTTGAGGACA	300		
2812	GCATGTGGAGATTGCACAGCGGACGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG	2871		
301	GCATGTGGAGATTGCACAGCGGACGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG	360		
2872	TGTGTGACCTCCAAATGCCCTATGTGGCCCTCCCTTCCCTTTGGCCAGTGTATGGAAATGGTAT	2931		
361	TGTGTGACCTCCAAATGCCCTATGTGGCCCTCCCTTCCCTTTGGCCAGTGTATGGAAATGGTAT	420		
2932	ACGATGAGCACTTGCCCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCATTGCTTG	2991		
421	ACGATGAGCACTTGCCCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCATTGCTTG	480		
2992	GAGCAACAGGCTGTGGCTGTGGTGTACTGATCCAGCAATCTGCGAAAGGGAATGCATA	3051		
481	GAGCAACAGGCTGTGGCTGTGGTGTACTGATCCAGCAATCTGCGAAAGGGAATGCATA	540		
3052	GAGGGTTCCCTATAAAGGACAGTGAAGATGCCCTTCGCAAGCCCTTACAGGAAATTTCTAT	3111		
541	GAGGGTTCCCTATAAAGGACAGTGAAGATGCCCTTCGCAAGCCCTTACAGGAAATTTCTAT	600		
3112	CCACAGCCCTTGCTCAATTCCAGCATGTGTTCTAGAGGACAGCAGATACAACCTGCTCTTTC	3171		
601	CCACAGCCCTTGCTCAATTCCAGCATGTGTTCTAGAGGACAGCAGATACAACCTGCTCTTTC	660		
3172	ATTCACTGTCCAGCTTGCCAA	3192		
661	ATTCACTGTCCAGCTTGCCAA	681		

RESULT 9
BM783788

```

LOCUS      BM783788      683 bp      mRNA      linear      EST 05-MAR-2002
DEFINITION K-EST00611794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
            mRNA sequence.
ACCESSION  BM783788
VERSION     BM783788.1 GI:19132020
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 683)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36 row: E column: 09
            High quality sequence stop: 683.
            Location/Qualifiers
                1..683
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S5SNU484-36-E09"
                /sex="M"
                /tissue_type="Stomach"
                /cell_type="Epithelial"
                /lab_host="Top10P"
                /clone_lib="S5SNU484"
                /note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI;
                Site: 2: NotI; The poly (A) + RNA was decapped with tobacco
                acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                including EcoRI site by treatment of T4 RNA ligase. The
                first strand cDNA was synthesized from oligo dT-selected
                mRNA by priming with dT-tailed vector. The dT-tailed
                vector was adjusted to have about 60nt. The cDNA vector
                was circularized with E. coli DNA ligase after digestion
                of EcoRI which site is also included in vector. An RNA
                strand converted to a DNA strand by Okayama-Berg method.
                The obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10P' by electroporation
                method."

ORIGIN
Query Match 15.9%; Score 680.4; DB 12; Length 683;
Best Local Similarity 99.7%; Pred. No. 3.1e-119;
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2512 CAGCTGGGAATATGCAAGTCAATCTCAGAGCATGTCGAAGTCACCTTAACCCCATGGTGC 2571
Db 1 CAGCTGGGAATATGCAAGTCAATCTCAGAGCATGTCGAAGTCACCTTAACCCCATGGTGC 60

QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCCCAATTACAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCCCAATTACAAT 120

QY 2632 AGTTTACTACAGTGGATCCGCTGTAGCCAGCATGCTGGATCTCTGTGGAATTTATCA 2691
Db 121 AGTTTACTACAGTGGATCCGCTGTAGCCAGCATGCTGGATCTCTGTGGAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCATCAATGTAGTGC 2751
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCATCAATGTAGTGC 240

QY 2752 TGTGAAGGCTCGAACAACACAGTGTACAGCATGTCGGAAGTCCGGAACCATGTGCTTGAGGACA 2811

```

```

Db 241 TGTGAAGGCTCGAACAACACAGTGTCTAAGCAGTGCCTGAGCAGTCCGGAACCATGTGCCCTTGAGGACA 300
QY 2812 GCATGTGGAGATTGCAACCGCGCAGCTCTGAGTGCATGTGGTGCAACATGAAGCAG 2871
Db 301 GCATGTGGAGATTGCAACCGCGCAGCTCTGAGTGCATGTGGTGCAACATGAAGCAG 360
QY 2872 TGTGTGAGCTCCAAATGCTGATGCGCTCTCCCTTTTGGCCAGTATGAAGTGTAT 2931
Db 361 TGTGTGAGCTCCAAATGCTGATGCGCTCTCCCTTTTGGCCAGTATGAAGTGTAT 420
QY 2932 ACCATGAGCAGCTGCCCCCTGAAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACCATGAGCAGCTGCCCCCTGAAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTG 480
QY 2992 GAGCAACAGGCTGTGGCTGTGTTACTGATCCAGCAATATCTGGCAAGGGAATGCATA 3051
Db 481 GAGCAACAGGCTGTGGCTGTGTTACTGATCCAGCAATATCTGGCAAGGGAATGCATA 540
QY 3052 GAGGGTTCTCTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 3111
Db 541 GAGGGTTCTCTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 600
QY 3112 CCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAATGCTTTTC 3171
Db 601 CCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAATGCTTTTC 660
QY 3172 ATTCACTGTCCAGCTTGCCAATG 3194
Db 661 ATTCACTGTCCAGCTTGCCAATG 693

RESULT 10
LOCUS    BM783258
DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
            mRNA sequence.
ACCESSION BM783258
VERSION    BM783258.1 GI:19131490
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
REFERENCE  1 (bases 1 to 680)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 31 row: H column: 08
            High quality sequence stop: 680.
            Location/Qualifiers
                1..680
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S5SNU484-31-H08"
                /sex="M"
                /tissue_type="Stomach"
                /cell_type="Epithelial"
                /cell_line="SNU-484"
                /lab_host="Top10P"
                /clone_lib="S5SNU484"
                /note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI;
                Site: 2: NotI; The poly (A) + RNA was decapped with tobacco
                acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                including EcoRI site by treatment of T4 RNA ligase. The
                first strand cDNA was synthesized from oligo dT-selected
                mRNA by priming with dT-tailed vector. The dT-tailed
                vector was adjusted to have about 60nt. The cDNA vector
                was circularized with E. coli DNA ligase after digestion
                of EcoRI which site is also included in vector. An RNA
                strand converted to a DNA strand by Okayama-Berg method.
                The obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10P' by electroporation
                method."

FEATURES
source

```


1. (bases 1 to 677)
 Oh,K.N.S., Hahn,Y.S., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Kim,Y.S., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: H column: 09
 High quality sequence stop: 677.
 Location/Qualifiers
 1. 677
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-36-H09"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /cell_line="SNU-484"
 /lab_host="Top10F"
 /clone_lib="S5SNU484"
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was decapped with tobacco
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
 including EcoRI site by treatment of T4 RNA ligase. The
 first strand cDNA was synthesized from oligo dt-selected
 mRNA by priming with dt-tailed vector. The dt-tailed
 vector was adjusted to have about 60nt. The cDNA vector
 was circularized with E. coli DNA ligase after digestion
 of EcoRI which site is also included in vector. An RNA
 strand converted to a DNA strand by Okayama-Berg method.
 The obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation
 method."

ORIGIN

Query Match 15.7%; Score 674.4; DB 12; Length 677;
 Best Local Similarity 99.7%; Pred. No. 4.3e-118; Indels 0; Gaps 0;
 Matches 675; Conservative 0; Mismatches 2;

Qy 2512 CAGCTGCGAAATAATGCAGTCATCTCAGAGCATGTCACAGCTCACTTAACCCCATGGGTC 2571
 Db 1 CAGCTGCGAAATAATGCAGTCATCTCAGAGCATGTCACAGCTCACTTAACCCCATGGGTC 60

Qy 2572 GGCCTTCGGAAGATCAATGTGTCCTATCTGGCTCTGGGAAGATATGTCCCATTTTACAAT 2631
 Db 61 GGCCTTCGGAAGATCAATGTGTCCTATCTGGCTCTGGGAAGATATGTCCCATTTTACAAT 120

Qy 2632 AGTTTACTACAGTGATGCCGCTCTGAGCCCATGATGCTGATCTCTGTGGAATTTATCA 2691
 Db 121 AGTTTACTACAGTGATGCCGCTCTGAGCCCATGATGCTGATCTCTGTGGAATTTATCA 180

Qy 2692 GAACCCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCCATCAATGGTAGTGTC 2751
 Db 181 GAACCCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCCATCAATGGTAGTGTC 240

Qy 2752 TGTGAAGGCTGCAAAACACAGTCTAAGCAGTCCCGGACACCATGTGCTTGAGGACA 2811
 Db 241 TGTGAAGGCTGCAAAACACAGTCTAAGCAGTCCCGGACACCATGTGCTTGAGGACA 300

Qy 2812 GCATGTGAGATTGCACCGCGGCAAGCTCTGAGTGCATGTGCTGCAGCAACATGAAGCAG 2871
 Db 301 GCATGTGAGATTGCACCGCGGCAAGCTCTGAGTGCATGTGCTGCAGCAACATGAAGCAG 360

Qy 2872 TGTGTGGACTCAATGCGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2933
 Db 361 TGTGTGGACTCAATGCGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

method."

ORIGIN

Query Match 15.7%; Score 673; DB 12; Length 673;
Best Local Similarity 100.0%; Pred.No. 7.9e-118;
Matches 673; Conservative 0; Mismatches 0; Indels

2512	QY	CAGCTGGCAATAATGCAGTGCATCTCAGAGCATGTCCAAGCTCACCTTTAAACCCCAATGGGTC	2571
1	Db	CAGCTGGCAATAATGCAGTGCATCTCAGAGCATGTCCAAGCTCACCTTTAAACCCCAATGGGTC	60
2572	QY	GGCCTTCGGAAGATCAATGTGTCTACTCGTGTGGGAAGATATGTCCCATTACAAAT	2631
61	Db	GGCCTTCGGAAGATCAATGTGTCTACTCGTGTGGGAAGATATGTCCCATTACAAAT	120
2632	QY	AGTTTACTACAGTGGATGCCGTCTGAGCCOAGTGAATGCTGGATCTGTGGAAATTTTATCA	2691
121	Db	AGTTTACTACAGTGGATGCCGTCTGAGCCOAGTGAATGCTGTGGAAATTTTATCA	180
2692	QY	GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTGTAGTGC	2751
181	Db	GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTGTAGTGC	240
2752	QY	TGTGAAGGCTGCAAAACCACTGCTTAAGCAGTCCGGACACCATGTGCTTTGAGGACA	2811
241	Db	TGTGAAGGCTGCAAAACCACTGCTTAAGCAGTCCGGACACCATGTGCTTTGAGGACA	300
2812	QY	GCATGTGGAGATTGCACAGCGGCAGCTCTGAGTGCATGTGGTGACGACACATGAGGCAG	2871
301	Db	GCATGTGGAGATTGCACAGCGGCAGCTCTGAGTGCATGTGGTGACGACACATGAGGCAG	360
2872	QY	TGTGTGGACTCCAATGCCTATGTGGGCTCCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT	2931
361	Db	TGTGTGGACTCCAATGCCTATGTGGGCTCCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT	420
2932	QY	ACGATGAGCACCTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTGTAGTCTATTGCTTG	2991
421	Db	ACGATGAGCACCTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTGTAGTCTATTGCTTG	480
2992	QY	GAGCAACGAGGCTGTGGCTGGTGTACTGTATCCAGCAATACTGGCAAGGAAATGCATA	3051
481	Db	GAGCAACGAGGCTGTGGCTGGTGTACTGTATCCAGCAATACTGGCAAGGAAATGCATA	540
3052	QY	GAGGGTTCTATAAAGGACGAGTGAAGTGCCTTTCGCAAGCCCCCTACAGGAAATTTCTAT	3111
541	Db	GAGGGTTCTATAAAGGACGAGTGAAGTGCCTTTCGCAAGCCCCCTACAGGAAATTTCTAT	600
3112	QY	CCACAGCCCCCTGCTCAATTCACAGCATGTGCTTAGGACAGCAGATPACAACTGGTCTTTTC	3171
601	Db	CCACAGCCCCCTGCTCAATTCACAGCATGTGCTTAGGACAGCAGATPACAACTGGTCTTTTC	660
3172	QY	ATTCACTGTCCAG	3184
661	Db	ATTCACTGTCCAG	673

RESULT 15

CD644584	LOCUS	792 bp	mrna	linear	EST 17-JUN-2003
CD644584	DEFINITION	AGENCOURT_1453938	NIA Human HI Embryonic Stem Cell	CDNA Library	
		(long)	Homo sapiens	cdna clone IMAGE:30421895	5', mRNA sequence.

ACCESSION CD644584

VERSION CD644584.1 GI:31816608

KEYWORDS EST.

SOURCE	Homo sapiens (human)
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

ORGANISM	Homo sapiens
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

I (bases 1 to 792)

NTH-MCC bttt. / mac sei ssk maw /
XUMHOB8

AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>

TITLE National Institutes of Health M

THE NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC) JOURNAL, Unpublished (1999)

COMMENT : Contact: Daniela S. Gerh
unpublished (1999)

[illegible]

ORIGIN

Query Match 15.5%; Score 663.8; DB 14; Length 792;
Best Local Similarity 94.9%; Pred. No. 4.8e-116;
Matches 686; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OV 479 GGCTCATTTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTCAATCATTTTGCTACAG 538

db
12 GGCTCATTTGAAGGACAGCCAAATAGAAATAATGAGACTTCGTTTCAATCATTTTGCTACAG 71

539 AGTGAGTTGGGACCA TTTATATGTTTATGATGGGGACTCAATTTATGCAACCGTAGTTG 598

db 72 AGTGAGTTGGGACCAATTTATATGTTTATGATGGGACATCAATTTATGCACCGGTAGTTG 131

QV 599 CTGCAATTAGTGGCCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACCTGTCCCTGAGGGTTG 658

132 CTGCAATTAGTGGCCCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACATGTTCCTGAGGTTG 191

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm1CA07 Bethesda, MD 20892
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Kc
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC c
one distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:

FEATURES

sources

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30421895"
 /issue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH103 (T1 phage-resistant)"
 /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Jlong)"
 /note=Vector: pCMV-Sport6; Site.1: NotI; Site.2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on DMF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MXG1, FUT3, SSEA-1, TUBB3, NES, GRAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off (with 5 ml pipette). RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-GACAGTCTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.3-4.9 of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Cetricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Cetricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb.

Qy 659 TTGCCACATCAGGTTATGCTTCTGCTCAATTTTTTTAGTGATGCTGCTTATAATTTGACTG 718
Db |||||||
192 TTGCCACATCAGGTTATGCTTCTGCTCAATTTTTTTAGTGATGCTGCTTATAATTTGACTG 251
Qy 719 GATTTAATATTACTTACAGTTTTCATATGTCCTCAATAAAGTCTCAGCCGAGGAGAGT 778
Db |||||||
252 GATTTAATATTACTTACAGTTTTCATATGTCCTCAATAAAGTCTCAGCCGAGGAGAGT 311
Qy 779 GTAAGATCAGTATAGCAGCGAAGTGTGAATGCTGATGCTTCTGAAAAGTCTGAAAAGGTG 838
Db |||||||
312 GTAAGATCAGTATAGCAGCGAAGTGTGAATGCTGATGCTTCTGAAAAGTCTGAAAAGGTG 371
Qy 839 AAGCATGTGACATTCCTCACTGACAGACAACTGTGTGTTTTCTCTCATCCAGGCAATCTGCA 898
Db |||||||
372 AAGCATGTGACATTCCTCACTGACAGACAACTGTGTGTTTTCTCTCATCCAGGCAATCTGCA 431
Qy 899 ATTCAAGTGTGACAGGATGCTCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTTCA 958
Db |||||||
432 ATTCAAGTGTGACAGGATGCTCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTTCA 491
Qy 959 TTCTGTACCAAGTAAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTAAAGCTCC 1018
Db |||||||
492 TTCTGTACCAAGTAAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTAAAGCTCC 551
Qy 1019 CCAGAGCATCTCAATAAAGCTGTGTCATATGCGAACAATTTATGCTGGTTGTTGGAGATATA 1078
Db |||||||
552 CCAGAGCATCTCAATAAAGCTGTGTCATATGCGAACAATTTATGCTGGTTGTTGGAGATATA 611
Qy 1079 TGTTCAACCACTCAGATTATAACATGCTTCTAGCGTATGACCTTCTTAGGGAGTGGC 1138
Db |||||||
612 TGTTCAACCACTCAGATTATAACATGCTTCTAGCGTATGACCTTCTTAGGGAGTGGC 671
Qy 1139 TTCCACTAAACCGTTCTGTGAACAATGTGTTGTAGATATGGTCAATCTTTGGCAATTAT 1198
Db |||||||
672 TTCACTAAACCGTTCTGTGAACAATGTGTTGTAGATATGGTCAATCTTTGGCAATTATACAGGAT 731
Qy 1199 ACA 1201
Db |||
732 AAA 734

Search completed: March 5, 2004, 12:07:59
Job time : 6801.26 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:06:34 ; Search time 1010.9 Seconds

(without alignments)
18028.269 Million cell updates/sec

Title: US-09-787-097-13

Sequence: 1 atggtggccagcagggggc.....agcctgggacctgcatctga 4290

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4290	100.0	4290	3	AAZ52276 Human mem
2	4290	100.0	8561	5	Aas72659 DNA encod
3	4184.8	97.5	8590	3	Aaz91920 Human mah
4	3981	92.8	4068	3	Aaz52275 Human mem
5	3801.4	88.6	3819	3	Aaz52277 Human sol
6	3696.2	86.2	4072	3	Aaz91921 Human mah
7	3536	82.4	8739	9	ADB53249 Primaty r
8	3469.4	80.9	3597	3	Aaz52274 Human sol
9	3469.4	80.9	3597	5	Aas72657 DNA encod
10	3469.4	80.9	3597	9	Add70997 Human att
11	3396.4	79.2	8927	3	Aaz91916 Wild type
12	2807.2	65.4	3490	5	Aas72660 DNA encod
13	2248.2	52.4	6255	3	Aaz91922 Human mah
14	1830.2	42.7	6373	3	Aaz91919 Human mah
15	1521.8	35.5	2419	3	Aaz91917 Murine ma
16	1479.4	34.5	4134	7	Abz58710 Murine ho
17	1467	34.2	4140	7	Abz58709 Human hom
18	1133.4	26.4	3609	6	Abz58709 Human cDN
19	873.2	20.4	3165	5	Aas78781 DNA encod
20	873.2	20.4	3313	5	Aas76294 DNA encod
21	819.4	19.1	1012	2	Aax84568 Biorhythm
22	812.8	18.9	6517	4	Aah98748 Human EST
23	773.2	18.0	6532	5	Aas86766 DNA encod

24	640.2	14.9	1051	3	AAZ91918	Aaz91918 Murine ma
25	420	9.8	1654	7	ABX34730	Abx34730 Human mdd
26	415.2	9.7	1383	6	AA154064	AA154064 cDNA of h
27	404.2	9.4	671	5	AAS72658	Aas72658 DNA encod
28	314.4	7.3	668	4	AAS02409	Aas02409 Human sec
29	277.4	6.5	481	4	AA14686	AA14686 Human bre
30	276.4	6.4	292	3	AAA45505	AAA45505 Human sec
31	267.6	6.2	845	6	ABQ52394	Abq52394 Oligonucl
32	267.6	6.2	845	6	ABQ52395	Abq52395 Oligonucl
33	266.4	6.2	482	4	AA123548	AA123548 Human bre
34	253.8	5.9	982	6	ABL39784	Ab139784 Human NS
35	249.2	5.8	531	4	AA143425	AA143425 Probe #12
36	249.2	5.8	531	4	ABS11523	ABs11523 Human gen
37	246	5.7	246	6	ABK38928	Abk38928 cDNA enco
38	246	5.7	246	6	ACA11257	ACA11257 Human lun
39	246	5.7	246	7	ACA02443	ACA02443 Lung "canc
40	236	5.5	657	7	ABZ58714	Abz58714 Upregulat
41	234.4	5.5	207433	5	ABZ72040	Abz72040 Gene 216
42	234.4	5.5	207433	7	ABX74891	ABx74891 BAC109812
43	209.8	4.9	5312	6	ABL32196	Ab132196 Human imm
44	206	4.8	845	6	ABQ52396	Abq52396 Oligonucl
45	206	4.8	845	6	ABQ52397	Abq52397 Oligonucl

ALIGNMENTS

RESULT 1
AAZ52276
ID AAZ52276 standard; cDNA; 4290 BP.
XX
AC AAZ52276;
DT 18-JUL-2000 (first entry)
XX
DE Human membrane attractin-2 cDNA.
XX
KW Human; membrane attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key
CDS Location/Qualifiers
FT 1..4290
FT /*tag= a
FT /product= "Membrane attractin-2"
XX
XX WC200015651-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-US020948.
XX
XX 14-SEP-1998; 98US-0100137P.
XX
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Duke-Cohan JS, Schlossman SF;
XX WPI; 2000-271373/23.
XX P-PSDB; AAY70691.
XX
XX Isolated nucleic acids encoding human attractin polypeptides useful for
XX enhancing immune responses.
XX
XX Claim 3; Fig 14; 120pp; English.
XX
XX The patent discloses four forms of human attractin polypeptides which
XX enhance immune response by promoting macrophage and monocyte spreading in
XX the presence of T cells. These include soluble attractin-1 and -2 and
XX membrane attractin-1 and -2. These various forms of attractin are encoded
XX by alternatively spliced mRNA molecule transcribed from a single gene.

QY 1921 CTGGTATTCACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTACAGCA 1980
Db 1921 CTGGTATTCACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTACAGCA 1980
QY 1981 GGAACCTGGTATTCGGTGTGTGGAAACAGAGGTGCTCTCAGTGTATCTCGTGGGCGGTG 2040
Db 1981 GGAACCTGGTATTCGGTGTGTGGAAACAGAGGTGCTCTCAGTGTATCTCGTGGGCGGTG 2040
QY 2041 GCAACTGATGAACAAGAAAGTGTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100
Db 2041 GCAACTGATGAACAAGAAAGTGTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100
QY 2101 CATGACAGATGTACACAGACACAGATGTTTACAGCTGTACAGCCAAACCACTGACTGC 2160
Db 2101 CATGACAGATGTACACAGACACAGATGTTTACAGCTGTACAGCCAAACCACTGACTGC 2160
QY 2161 CACTGSGCAATGACCAATGTTGTCCTCCAGGAAACACAGCTGTCTCAGAGCCGAGATCTCC 2220
Db 2161 CACTGSGCAATGACCAATGTTGTCCTCCAGGAAACACAGCTGTCTCAGAGCCGAGATCTCC 2220
QY 2221 ATTTTATAGTATGAGAAATGCCCCCAGGATAACCCCATGTACTACTGTACACAGAGACC 2280
Db 2221 ATTTTATAGTATGAGAAATGCCCCCAGGATAACCCCATGTACTACTGTACACAGAGACC 2280
QY 2281 AGCTGACAGAGCTGTGCTCGGACCAAGTGCAGTGGGAGCCCGGAAATCAGGAGTGC 2340
Db 2281 AGCTGACAGAGCTGTGCTCGGACCAAGTGCAGTGGGAGCCCGGAAATCAGGAGTGC 2340
QY 2341 ATTGCCCCCGGAAATATCTGTGGCAATGCTGGCAATGCTGGCAATGCTGGCAATGCTGGCA 2400
Db 2341 ATTGCCCCCGGAAATATCTGTGGCAATGCTGGCAATGCTGGCAATGCTGGCAATGCTGGCA 2400
QY 2401 AAAATTAATCTGCGAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAACAT 2460
Db 2401 AAAATTAATCTGCGAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAACAT 2460
QY 2461 GCCCTTTTGGCTTCTCTTACACCCAGGAAAGTGAATTTGTCCTTAAGCAGTGGGA 2520
Db 2461 GCCCTTTTGGCTTCTCTTACACCCAGGAAAGTGAATTTGTCCTTAAGCAGTGGGA 2520
QY 2521 ATAAAGCAGTCACTCAGAGCATGTCCAAGTCACTTAACCCCATGGGTGCGCTCGG 2580
Db 2521 ATAAAGCAGTCACTCAGAGCATGTCCAAGTCACTTAACCCCATGGGTGCGCTCGG 2580
QY 2581 AAGATCAATGTGCTACTGTGCGGAAAGATATGTCCCAATTTACAAATAGTTTACTA 2640
Db 2581 AAGATCAATGTGCTACTGTGCGGAAAGATATGTCCCAATTTACAAATAGTTTACTA 2640
QY 2641 CAGTGGATGCGCTCTGAGCCAGTGTGCTGGATCTGTGGAAATTTATCAGAACCCAGT 2700
Db 2641 CAGTGGATGCGCTCTGAGCCAGTGTGCTGGATCTGTGGAAATTTATCAGAACCCAGT 2700
QY 2701 ACTGGGAGCTGAAGGCTGCAACCTGCATCAACCACTCAATGTGTAGTGTCTGTGAAAGG 2760
Db 2701 ACTGGGAGCTGAAGGCTGCAACCTGCATCAACCACTCAATGTGTAGTGTCTGTGAAAGG 2760
QY 2761 CCTGCAAAACCAAGTGTGCTGAGTGGCGGACACCAATGTGCTGTGAGGACAGCATGTGGA 2820
Db 2761 CCTGCAAAACCAAGTGTGCTGAGTGGCGGACACCAATGTGCTGTGAGGACAGCATGTGGA 2820
QY 2821 GATTGACACAGGAGCTCTGAGTGCATGTGTGTGAGACACATGAGCAGTGTGTGGAC 2880
Db 2821 GATTGACACAGGAGCTCTGAGTGCATGTGTGTGAGACACATGAGCAGTGTGTGGAC 2880
QY 2881 TCCAAATGCTATGTGGCT 2940
Db 2881 TCCAAATGCTATGTGGCT 2940
QY 2941 ACCTGCCCCCTGAAATGTTTACGCTACTGTACCTGTAGTCAATGCTGGAGCAACCA 3000
Db 2941 ACCTGCCCCCTGAAATGTTTACGCTACTGTACCTGTAGTCAATGCTGGAGCAACCA 3000

QY 3001 GGCTGTGGCTGTACTGATCCAGCAATACCTGGCAAGGAAATGCATAGAGGTTCC 3060
Db 3001 GGCTGTGGCTGTACTGATCCAGCAATACCTGGCAAGGAAATGCATAGAGGTTCC 3060
QY 3061 TATAAAGAACAGTGAAGATGCCCTTCGAAAGCCCTACAGGAAATTTCTATCCACAGCCC 3120
Db 3061 TATAAAGAACAGTGAAGATGCCCTTCGAAAGCCCTACAGGAAATTTCTATCCACAGCCC 3120
QY 3121 CTGCTCAATTCACCATGTGTCTAGAGACAGAGATCAACCTGCTCTTTCAATCACTGT 3180
Db 3121 CTGCTCAATTCACCATGTGTCTAGAGACAGAGATCAACCTGCTCTTTCAATCACTGT 3180
QY 3181 CCAGCTTGGCAATGCAACGCGCCACAGTAAATGCAATCAGAGCATCTGTGAGAGTGT 3240
Db 3181 CCAGCTTGGCAATGCAACGCGCCACAGTAAATGCAATCAGAGCATCTGTGAGAGTGT 3240
QY 3241 GAGAACCTGACCAAGGCAAGCTGCGAGACCTGCAATCTGCTTCTACGCTGATCCC 3300
Db 3241 GAGAACCTGACCAAGGCAAGCTGCGAGACCTGCAATCTGCTTCTACGCTGATCCC 3300
QY 3301 ACCAATGAGGAGAAATGTCAAGTGCATGCAATGGGCAAGCTCTCTGTGCAACACC 3360
Db 3301 ACCAATGAGGAGAAATGTCAAGTGCATGCAATGGGCAAGCTCTCTGTGCAACACC 3360
QY 3361 AACAGGCAAGTGTCTTCTGACCAACCAAGGCGTCAAGGGGAGGAGTCCAGCTATGT 3420
Db 3361 AACAGGCAAGTGTCTTCTGACCAACCAAGGCGTCAAGGGGAGGAGTCCAGCTATGT 3420
QY 3421 GAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTAATCTTCTT 3480
Db 3421 GAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTAATCTTCTT 3480
QY 3481 ATGTACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTAACAGCTATC 3540
Db 3481 ATGTACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTAACAGCTATC 3540
QY 3541 AATTTTGTGGTACTCTTCGAGCAACAAAGGAAATTTGGACATGTTTCAATGCCCTCC 3600
Db 3541 AATTTTGTGGTACTCTTCGAGCAACAAAGGAAATTTGGACATGTTTCAATGCCCTCC 3600
QY 3601 AAGAAATTTCAACCTCAACATCACTGGCTGGCTGCCAGTTCCTCAGCTGGAACCCAGGCTGGA 3660
Db 3601 AAGAAATTTCAACCTCAACATCACTGGCTGGCTGCCAGTTCCTCAGCTGGAACCCAGGCTGGA 3660
QY 3661 GAAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
Db 3661 GAAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
QY 3721 GAGAAATTTGATTTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAAGTAAATTC 3780
Db 3721 GAGAAATTTGATTTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAAGTAAATTC 3780
QY 3781 ACTGSCCCATCAAAATTCAGATTCGCTTCTCAGCAGCAAGCAATTTTATGCAAGCTGTA 3840
Db 3781 ACTGSCCCATCAAAATTCAGATTCGCTTCTCAGCAGCAAGCAATTTTATGCAAGCTGTA 3840
QY 3841 CAGTCTTCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
Db 3841 CAGTCTTCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
QY 3901 TGAAGATCAAAACAAAGTGTGTTGGGCTCCAGACCTGAGAGCAACCTTCTTGAGAGATG 3960
Db 3901 TGAAGATCAAAACAAAGTGTGTTGGGCTCCAGACCTGAGAGCAACCTTCTTGAGAGATG 3960
QY 3961 CAACAGATGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Db 3961 CAACAGATGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
QY 4021 CCTCTCTGATCTTATTTGGGGGAGTAAAGACTGTTCCAAACCCCAATTCATCTGAGCGC 4080
Db 4021 CCTCTCTGATCTTATTTGGGGGAGTAAAGACTGTTCCAAACCCCAATTCATCTGAGCGC 4080
QY 4081 TGTTTTGGCAACAAAGCCGCTGCTCTCTGTTGTTGAGGCTCCTCTGAGGCTGCTGCTGCTGCT 4140

D _b	4 081	TGTTTTGGCAACAAGCGGTGTCCTCTCTGTGTTGTGAGGCTCCCTCAGGCGCTGGGT	4 140
Q _y	4 141	GGCATCCTCTCCTCTGGGCAGTCAGGTCCTTGCTGTGGCCAGCGCCCTGTGTGGACATTCT	4 200
D _b	4 141	GGCATCCTCTCCTCTGGGCAGTCAGGTCCTTGCTGTGGCCAGCGCCCTGTGTGGACATTCT	4 200
Q _y	4 201	CAGCAGATCGCGATAGTGTACAGGAAAGTCAGGAGCCCTTGAGAAACCGGAGACGACG	4 260
D _b	4 201	CAGCAGATCGCGATAGTGTACAGGAAAGTCAGGAGCCCTTGAGAAACCGGAGACGACG	4 260
Q _y	4 261	CCCCCTGCAGCCTGGAGCCTGCATCTGA	4 290
D _b	4 261	CCCCCTGCAGCCTGGAGCCTGCATCTGA	4 290

RESULT 2

AAS72659

ID AAS72659 standard; cDNA; 8561 BP.



AC AAS72659;

XX
DT 13-FEB-2002 (first entry)

XX
DE DNA encoding novel human diagnostic protein #8463.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder;

OS Homo sapiens.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PN .WO200175067-A2.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PD 11-OCT-2001.

XX
DE
30-MAR-2001: 2001W

30-MAR-2001; 2001WU-US008631.
FF
XX

31-MAR-2000: 2000US-

PR 23-AUG-2000: 2000

23-A03-2000, 200003-00049167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX

PI Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.
P-PSDB; ABG08472.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 8463; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in

QY 961 CCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATATCTTAACCTTAAAGCTCCCC 1020
Db 961 CCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATATCTTAACCTTAAAGCTCCCC 1020
QY 1021 AGAGCATCTCATAAAGCTGGTCAATGGAAGAAATATGATGGGTGTGTGGAGGATATG 1080
Db 1021 AGAGCATCTCATAAAGCTGGTCAATGGAAGAAATATGATGGGTGTGTGGAGGATATG 1080
QY 1081 TTCACACACTCAGATTATAACATGGTCTAGCGTATGACCTTGTCTTACGGAGTGGCTT 1140
Db 1081 TTCACACACTCAGATTATAACATGGTCTAGCGTATGACCTTGTCTTACGGAGTGGCTT 1140
QY 1141 CCACATAACCGTCTGTGAAACAATGTGGTGTGTAGATPATGGTCAATCTTTGGCAATATAC 1200
Db 1141 CCACATAACCGTCTGTGAAACAATGTGGTGTGTAGATPATGGTCAATCTTTGGCAATATAC 1200
QY 1201 AAGGATAAATTTACATGTATGGAGGAATATGATCCAACTGGGATGTGACCAATGAG 1260
Db 1201 AAGGATAAATTTACATGTATGGAGGAATATGATCCAACTGGGATGTGACCAATGAG 1260
QY 1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTGTGACCCCTAAGGCAAGGAG 1320
Db 1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTGTGACCCCTAAGGCAAGGAG 1320
QY 1321 CAGTATGAGTGGTGGGCACTCTGCACACATTTGTTACACTGAAGAATGSCCGAGTGGTC 1380
Db 1321 CAGTATGAGTGGTGGGCACTCTGCACACATTTGTTACACTGAAGAATGSCCGAGTGGTC 1380
QY 1381 ATGCTGGTCACTTTTGTGCTACCTGCTCTCATGATATATAAGCAATGTGCAAGGAATAT 1440
Db 1381 ATGCTGGTCACTTTTGTGCTACCTGCTCTCATGATATATAAGCAATGTGCAAGGAATAT 1440
QY 1441 GATTGGGATAAGAACACATGAGTATATTAACACCCAGGGTGCCTTGTGCAAGGGGT 1500
Db 1441 GATTGGGATAAGAACACATGAGTATATTAACACCCAGGGTGCCTTGTGCAAGGGGT 1500
QY 1501 TACGGCCATACAGTGTGTAGACATAGGACAGGCGCCCTATAGCTTCATGTTGGCTAC 1560
Db 1501 TACGGCCATACAGTGTGTAGACATAGGACAGGCGCCCTATAGCTTCATGTTGGCTAC 1560
QY 1561 AAGGCTTTCAGTGCCCAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAT 1620
Db 1561 AAGGCTTTCAGTGCCCAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAT 1620
QY 1621 ACCGAGTGTGGACCATCTTTAAGGACAGCGGATTTTCGGTACTTGTGCACAGCTGTG 1680
Db 1621 ACCGAGTGTGGACCATCTTTAAGGACAGCGGATTTTCGGTACTTGTGCACAGCTGTG 1680
QY 1681 ATAGTGTGTGAAACCATGCTGTGTGTGGGGAACACACACATGACATCTATGAGC 1740
Db 1681 ATAGTGTGTGAAACCATGCTGTGTGTGGGGAACACACACATGACATCTATGAGC 1740
QY 1741 CATGGCGCAATGCTTCTTTCAGATTTCAITGGCTATGACATGCTGTGACCGCTGG 1800
Db 1741 CATGGCGCAATGCTTCTTTCAGATTTCAITGGCTATGACATGCTGTGACCGCTGG 1800
QY 1801 TCAGTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Db 1801 TCAGTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
QY 1861 TTACACACAGCACCATGATGTGTGTGGTGTTCATAGTCTCTCTCCCTCAGCGACATC 1920
Db 1861 TTACACACAGCACCATGATGTGTGTGGTGTTCATAGTCTCTCTCTCAGCGACATC 1920
QY 1921 CTGGTATTCACTCGGAACAGTGTGATGGGATCGAGTGAAGCGGCTTGTATTAGCAGCA 1980
Db 1921 CTGGTATTCACTCGGAACAGTGTGATGGGATCGAGTGAAGCGGCTTGTATTAGCAGCA 1980
QY 1981 GGACCTGGTATTCTGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTG 2040
Db 1981 GGACCTGGTATTCTGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTG 2040
QY 2041 GCAACTGATGAACAAAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100

Db 2041 GCAACTGATGAACAAAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100
QY 2101 CATGACAGATGACACAGACACAGATTGTACAGCTGTACAGCCAAACCAATGACTGTC 2160
Db 2101 CATGACAGATGACACAGACACAGATTGTACAGCTGTACAGCCAAACCAATGACTGTC 2160
QY 2161 CACTGGTCAATGACCAATTTGTTCCCGAGGAACCAACAGCTGTCTCAGAAAGCCAGATCTCC 2220
Db 2161 CACTGGTCAATGACCAATTTGTTCCCGAGGAACCAACAGCTGTCTCAGAAAGCCAGATCTCC 2220
QY 2221 ATTTTGTAGTATGAGATTTGCCCAAGGATATACCAATGCTTACTGTATAAAGAAAGACC 2280
Db 2221 ATTTTGTAGTATGAGATTTGCCCAAGGATATACCAATGCTTACTGTATAAAGAAAGACC 2280
QY 2281 AGCTGACAGAGCTGTGCTCCCTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2340
Db 2281 AGCTGACAGAGCTGTGCTCCCTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2340
QY 2341 ATTGCCCTCCCGAAAAATCTGTGGCATTTGGTGGCAATTTGGTGGAAACTCATGTTTG 2400
Db 2341 ATTGCCCTCCCGAAAAATCTGTGGCATTTGGTGGCAATTTGGTGGAAACTCATGTTTG 2400
QY 2401 AAAATTTACTTACGCAAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAACAT 2460
Db 2401 AAAATTTACTTACGCAAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAACAT 2460
QY 2461 GGCCTTTTGGCTTCTTAAACCCAGAAAGAGTAGAATTTGCTTAAAGCAGCTGCGA 2520
Db 2461 GGCCTTTTGGCTTCTTAAACCCAGAAAGAGTAGAATTTGCTTAAAGCAGCTGCGA 2520
QY 2521 ATAAAGCAGTCACTCAGAGCATGTCACAGCTCACTTAAACCCATGGGTGGGCTTGG 2580
Db 2521 ATAAAGCAGTCACTCAGAGCATGTCACAGCTCACTTAAACCCATGGGTGGGCTTGG 2580
QY 2581 ATAAAGCAGTCACTCAGAGCATGTCACAGCTCACTTAAACCCATGGGTGGGCTTGG 2580
Db 2581 ATAAAGCAGTCACTCAGAGCATGTCACAGCTCACTTAAACCCATGGGTGGGCTTGG 2580
QY 2581 AAGATCAATGTGTCTTACTTGGTGGGAGGATATGTCCTTCAATTTTCAAAATGTTTACTA 2640
Db 2581 AAGATCAATGTGTCTTACTTGGTGGGAGGATATGTCCTTCAATTTTCAAAATGTTTACTA 2640
QY 2641 CAGTGAATGCCCTGTGAGCCAGTGTGATGCTGTGGAATTTTATCAGAACCCAGT 2700
Db 2641 CAGTGAATGCCCTGTGAGCCAGTGTGATGCTGTGGAATTTTATCAGAACCCAGT 2700
QY 2701 ACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCACTCAATGGTAGTCTGTGAAAGG 2760
Db 2701 ACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCACTCAATGGTAGTCTGTGAAAGG 2760
QY 2761 CTTGCAAAACACAGTGTCTAAGCAGTCCCGGACACATGTGCCTTGAGGACAGCATGTGA 2820
Db 2761 CTTGCAAAACACAGTGTCTAAGCAGTCCCGGACACATGTGCCTTGAGGACAGCATGTGA 2820
QY 2821 GATTGACACAGCGGAGCTGTGATGTCATGTGTGCGCAACATGAAGCAGTGTGTGGAC 2880
Db 2821 GATTGACACAGCGGAGCTGTGATGTCATGTGTGCGCAACATGAAGCAGTGTGTGGAC 2880
QY 2881 TCCAATGCCCTATGTGGCTCTTTCCTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Db 2881 TCCAATGCCCTATGTGGCTCTTTCCTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
QY 2941 ACCTGCCCTTGAATTTGTTTTCAGCTACTGTACCTGTAGTCACTTGTGGAGCAACCA 3000
Db 2941 ACCTGCCCTTGAATTTGTTTTCAGCTACTGTACCTGTAGTCACTTGTGGAGCAACCA 3000
QY 3001 GGCTGTGGCTGTGTACTGTATCCAGCAATACTTGGCAAAAGGGAATGCAATAGAGGTTC 3060
Db 3001 GGCTGTGGCTGTGTACTGTATCCAGCAATACTTGGCAAAAGGGAATGCAATAGAGGTTC 3060
QY 3061 TATAAAGGACCAAGTGAAGTGCCTTGGAGCCCTTACAGGAATTTCTATCCACAGCC 3120
Db 3061 TATAAAGGACCAAGTGAAGTGCCTTGGAGCCCTTACAGGAATTTCTATCCACAGCC 3120
QY 3121 CTGCTCAATTCACAGCATGTCTGTAGAGGACAGCAGATACAACTGGTCTTCTTCACTGT 3180

Db 3121 CTGCTCAATTCAGCATGTGCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCACTGT 3180
QY 3181 CCAGCTTGCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAGTGT 3240
Db 3181 CCAGCTTGCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAGTGT 3240
QY 3241 GAGAAGCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
Db 3241 GAGAAGCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
QY 3301 ACAAATGGAGGAATGTGAGCCATGCAAGTGCATGCGCAAGCGCTCTCTGTGCAACACC 3360
Db 3301 ACCAAATGGAGGAATGTGAGCCATGCAAGTGCATGCGCAAGCGCTCTCTGTGCAACACC 3360
QY 3361 AACACGGGCAAGTCTCTGCAACCAACCAAGGCGCTCAAGGGGAGAGTGCAGCTATGT 3420
Db 3361 AACACGGGCAAGTCTCTGCAACCAACCAAGGCGCTCAAGGGGAGAGTGCAGCTATGT 3420
QY 3421 GAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTA TACTCTCTT 3480
Db 3421 GAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTA TACTCTCTT 3480
QY 3481 ATTGACTATCAGTTTCACTTTAGTCTATCCAGGAGATGATCGTATTACACAGCTATC 3540
Db 3481 ATTGACTATCAGTTTCACTTTAGTCTATCCAGGAGATGATCGTATTACACAGCTATC 3540
QY 3541 AATTTTGTGGTACTCTCTGACGAACAAACAGGGAATTTGGACATGTTTCATCAATGCCCTCC 3600
Db 3541 AATTTTGTGGTACTCTCTGACGAACAAACAGGGAATTTGGACATGTTTCATCAATGCCCTCC 3600
QY 3601 AAGAAATTCACCTCAACATCACCTGGGCTGCGAGTTCTCAGCTGGAACCCAGGCTGGA 3660
Db 3601 AAGAAATTCACCTCAACATCACCTGGGCTGCGAGTTCTCAGCTGGAACCCAGGCTGGA 3660
QY 3661 GAAGAGATGCTGTGTTTCAAAACCAACCAATTAAGGAGTACAAAGATAGTTCTCTTAAT 3720
Db 3661 GAAGAGATGCTGTGTTTCAAAACCAACCAATTAAGGAGTACAAAGATAGTTCTCTTAAT 3720
QY 3721 GAGAAATTTGATTTTGGCAACCCCAATATCACTTTCTTTGTTTATGTCAGTAATTC 3780
Db 3721 GAGAAATTTGATTTTGGCAACCCCAATATCACTTTCTTTGTTTATGTCAGTAATTC 3780
QY 3781 ACCTGGCCCAATCAAAATTCAGATGCTCTCTCAGCAGCAATTTATGAGCTGGTA 3840
Db 3781 ACCTGGCCCAATCAAAATTCAGATGCTCTCTCAGCAGCAATTTATGAGCTGGTA 3840
QY 3841 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
Db 3841 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
QY 3901 TGGAGATCAAAAGTGTGGGCTGCGAGCTGAGAGCAATCTTCTTCAGAGATG 3960
Db 3901 TGGAGATCAAAAGTGTGGGCTGCGAGCTGAGAGCAATCTTCTTCAGAGATG 3960
QY 3961 CAACAGATGGCCAGCGCT 4020
Db 3961 CAACAGATGGCCAGCGCT 4020
QY 4021 CCTCTGATCTTATTTGGGGGAGTATAAGACTTTTCCAAACCCATTTGCACTGGAGCCG 4080
Db 4021 CCTCTGATCTTATTTGGGGGAGTATAAGACTTTTCCAAACCCATTTGCACTGGAGCCG 4080
QY 4081 TGTTTTGGCAACAAAGCGCT 4140
Db 4081 TGTTTTGGCAACAAAGCGCT 4140
QY 4141 GGCAATCCCTCTCTCTGGGAGTACAGTCTTCTCTGTCGCGCCCTCTGTCGACATTTCT 4200
Db 4141 GGCAATCCCTCTCTCTGGGAGTACAGTCTTCTCTGTCGCGCCCTCTGTCGACATTTCT 4200
QY 4201 CAGCAGATGCGGATGTTACAAAGAGAGTACAGGAGCTGAGAAACCCGAGAGCAGCAG 4260
Db 4201 CAGCAGATGCGGATGTTACAAAGAGAGTACAGGAGCTGAGAAACCCGAGAGCAGCAG 4260

QY 4261 CCCCTGCAAGCTGGGACCTGCACTCTGA 4290
Db 4261 CCCCTGCAAGCTGGGACCTGCACTCTGA 4290

RESULT 3

AAZ91920
ID AAZ91920 standard; cDNA; 8590 BP.

XX AAZ91920;

XX 08-JUN-2000 (first entry)

XX Human mahogany protein coding sequence #2.

XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antioesity; antianorexic; anticachexic; ss.

OS Homo sapiens.

XX WO200005373-A2.

XX 03-FEB-2000.

XX 21-JUL-1999; 99WO-US016484.

XX 21-JUL-1998; 98US-0093630P.

XX 20-OCT-1998; 98US-0104978P.

XX 05-FEB-1999; 99US-00245041.

XX (MILL-) MILLENIUM PHARM INC.

XX Moore K, Nagle DL;

XX WPI; 2000-195103/17.

XX P-PSDB; AAY81807.

XX New human and murine mahogany genes, useful, e.g. for diagnosis and

XX treatment of body weight disorders.

XX Claim 1; Fig 18a; 188pp; English.

XX This sequence represents a human mahogany gene of the invention. The
mahogany genes are used: (i) to produce recombinant mahogany (mg)
proteins (ii); (ii) as a source of antisense, ribozyme or triplex-forming
therapeutics; (iii) as a source of diagnostic probes and primers for
detecting expression of mg genes or mutations, regulatory defects, in
this gene, or for isolation of related sequences; and (iv) in (cell-
based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
identify other (extra)cellular products involved in weight regulation,
and to screen for agents that disrupt interaction between (ii) and other
macromolecules. The Ab are used to detect abnormal levels (or function)
of (ii) (for diagnosis, prognosis or monitoring of treatment); to
evaluate (ii)-expressing cells intended for cell therapy, and as
therapeutic mg inhibitors. Cells that express the mg gene (or contain the
mg polypeptide) are used to identify agents (A) that modulate mg
activity. (A) are potentially useful for the treatment of body weight
disorders, particularly obesity, cachexia or anorexia, or other
conditions associated with the mg gene such as hyperpigmentation,
hyperphagia and disorders that result in increased metabolic rate

XX Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

Query Match 97.5%; Score 4184.8; DB 3; Length 8590;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 4232; Conservative 0; Mismatches 47; Indels 9; Gaps 1;

QY 3 GGTGCCCCAGCGCGCAACTGAGGCAAGCTGAGGAGGAGGAGCGCGGCGAGCGGAGC 62

Db 9 GGTGCCCCAGCGCGCGCAACTGAGGCAAGCTGAGGAGGAGGAGCGCGGCGAGCGGAGC 68

Db 2220 TTTTAGGTATGAGAAATGCCCCAGGATAAACCTATGTACTGTAAACAAGAACCCAG 2279
Qy 2283 CTGAGGAGCTGTGCCCTGGACCAACTGCCAGTGGGAGCCCGGGAATCAGGAGTGCAT 2342
Db 2280 CTGAGGAGCTGTGCCCTGGACCAACTGCCAGTGGGAGCCCGGGAATCAGGAGTGCAT 2339
Qy 2343 TCCCTTGCCGCAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGAACATCATGTTTGA 2402
Db 2340 TCCCTTGCCGCAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGAACATCATGTTTGA 2399
Qy 2403 AATTACTCTGCAAGGAGAAATATGCAATGCTTAATGCTTCTGTAGGACCAATGC 2462
Db 2400 AATTACTCTGCAAGGAGAAATATGCAATGCTTAATGCTTCTGTAGGACCAATGC 2459
Qy 2463 CTTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTCTTAAGCAGCTGCCAAT 2522
Db 2460 CTTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTCTTAAGCAGCTGCCAAT 2519
Qy 2523 AATGAGTATCTCAGAGCATGTCCAGCTCAGCTTAACCCATGGGTGGCTTCGGAA 2582
Db 2520 AATGAGTATCTCAGAGCATGTCCAGCTCAGCTTAACCCATGGGTGGCTTCGGAA 2579
Qy 2583 GATCAATGTGCTTACTGCTGCGAAGATATGTCCCAATTTACAATAGTTTACTACA 2642
Db 2580 GATCAATGTGCTTACTGCTGCGAAGATATGTCCCAATTTACAATAGTTTACTACA 2639
Qy 2643 GTGGATGCCGTCTGAGCCAGTGTGGTGGATTTCTGTGGAAATTTATCAGAACCCAGTAC 2702
Db 2640 GTGGATGCCGTCTGAGCCAGTGTGGTGGATTTCTGTGGAAATTTATCAGAACCCAGTAC 2699
Qy 2703 TCGGGGATGAGAGCTGCAACTGCATCAACCACTCAATGGTAGTCTGTGAAAGGCC 2762
Db 2700 TCGGGGATGAGAGCTGCAACTGCATCAACCACTCAATGGTAGTCTGTGAAAGGCC 2759
Qy 2763 TCGAAACACAGTGTGAAGCTGCTGCGGACACCACTGTGCGCTTGAGGACAGCATGTGGAGA 2822
Db 2760 TCGAAACACAGTGTGAAGCTGCTGCGGACACCACTGTGCGCTTGAGGACAGCATGTGGAGA 2819
Qy 2823 TTGCACAGCGGAGCTGTGAGTGCATGTGTGAGCAACATGAGCAGTGTGTGAGCTC 2882
Db 2820 TTGCACAGCGGAGCTGTGAGTGCATGTGTGAGCAACATGAGCAGTGTGTGAGCTC 2879
Qy 2883 CAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAGCAC 2942
Db 2880 CAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAGCAC 2939
Qy 2943 CTGCCCCCTGAAATTTGTTTGGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCAACGAGG 3002
Db 2940 CTGCCCCCTGAAATTTGTTTGGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCAACGAGG 2999
Qy 3003 CTGTGGCTGTGTACTGTATCCAGCAATCTGCAAGGAAATGCAATGAGAGGTTCCCTA 3062
Db 3000 CTGTGGCTGTGTACTGTATCCAGCAATCTGCAAGGAAATGCAATGAGAGGTTCCCTA 3059
Qy 3063 TAAAGGACCAAGTGAAGTGTCTGCAAGCCCTTACAGGAAATTTCTATCCAGGCCCT 3122
Db 3060 TAAAGGACCAAGTGAAGTGTCTGCAAGCCCTTACAGGAAATTTCTATCCAGGCCCT 3119
Qy 3123 GCTCAATTCAGCATGTGTCTAGAGGACAGCATCAATGCTGCTTCTTCACTGTGC 3182
Db 3120 GCTCAATTCAGCATGTGTCTAGAGGACAGCATCAATGCTGCTTCTTCACTGTGC 3179
Qy 3183 AGCTTGCAATGCAAGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242
Db 3180 AGCTTGCAATGCAAGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3239
Qy 3243 GAACCTGACACAGGCAAGCACTGCGAGACCTGTGATCTGTGCTTCTACGGTATCCAC 3302
Db 3240 GAACCTGACACAGGCAAGCACTGCGAGACCTGTGATCTGTGCTTCTACGGTATCCAC 3299
Qy 3303 CAATGAGGAAATGTCTAGGCATCAATGCAATGCGACGCTCTCTGTGTGCAACCA 3362

Db 3300 CAATGAGGAAATGTCTAGGCATCAATGGCAACGCTCTCTGTGTCAACACCA 3359
Qy 3363 CACGGCAAGTGTCTCTGACCAACCAAGGCGCTCAGGGGAGAGTGCACGCTATGTGA 3422
Db 3360 CACGGCAAGTGTCTCTGACCAACCAAGGCGCTCAGGGGAGAGTGCACGCTATGTGA 3419
Qy 3423 GGTAGAAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTACTCTTCTTAT 3482
Db 3420 GGTAGAAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTACTCTTCTTAT 3479
Qy 3483 TGAATATCATGTTTCACTTCTATCCAGAGAGTATGATGCTATTTACAGAGCTATCAA 3542
Db 3480 TGAATATCATGTTTCACTTCTATCCAGAGAGTATGATGCTATTTACAGAGCTATCAA 3539
Qy 3543 TTTTGTGGCTACTCTCTGACGAAACCAACAGGATTTGGACATGTTTCATCAATGCCCTCAA 3602
Db 3540 TTTTGTGGCTACTCTCTGACGAAACCAACAGGATTTGGACATGTTTCATCAATGCCCTCAA 3599
Qy 3603 GAATTTCAACCTCAACATCACTGGGCTGCAGTTCCTCAGCTGGAAACCCAGGCTGGAGA 3662
Db 3600 GAATTTCAACCTCAACATCACTGGGCTGCAGTTCCTCAGCTGGAAACCCAGGCTGGAGA 3659
Qy 3663 AGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAAATGA 3722
Db 3660 AGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAAATGA 3719
Qy 3723 GAATTTGATTTTTCGCAACCCAAATATCACTTCTTGTGTTTATGTCAAGTAATTTTCA 3782
Db 3720 GAATTTGATTTTTCGCAACCCAAATATCACTTCTTGTGTTTATGTCAAGTAATTTTCA 3779
Qy 3783 CTGGCCCAATCAAAATTCAGATTTGCTCTCTCAGCAGACAGCAATTTATGAGCTGGTACA 3842
Db 3780 CTGGCCCAATCAAAATTCAGATTTGCTCTCTCAGCAGACAGCAATTTATGAGCTGGTACA 3839
Qy 3843 GTTCTTCTGAGCTTCTTCTCAGTGTGTTTCTCTCTCTTCTGCTGCTGCTGCTGCTGCTG 3902
Db 3840 GTTCTTCTGAGCTTCTTCTCAGTGTGTTTCTCTCTCTTCTGCTGCTGCTGCTGCTGCTG 3899
Qy 3903 GAAGATCAAAACAAAGTGTGTTGGCTCCAGAGCTAGAGAGCACTTCTTCAGAGATGCA 3962
Db 3900 GAAGATCAAAACAAAGTGTGTTGGCTCCAGAGCTAGAGAGCACTTCTTCAGAGATGCA 3959
Qy 3963 ACAGATGCGCAGCGCTGCTTCTGCTCTGTAATGCTGCTTGGAAACAGATGAGGAGCC 4022
Db 3960 ACAGATGCGCAGCGCTGCTTCTGCTCTGTAATGCTGCTTGGAAACAGATGAGGAGCC 4019
Qy 4023 TCCTGATCTTATTTGGGGGAGTATTAAGACTGTGTTCCCAAAACCAATGCACTGGAGCCGTG 4082
Db 4020 TCCTGATCTTATTTGGGGGAGTATTAAGACTGTGTTCCCAAAACCAATGCACTGGAGCCGTG 4079
Qy 4083 TTTTGGCAACAAAGCGCTGCTCTCTGCTGTTTGTGAGGCTCCCTCGAGGCTGGGTGG 4142
Db 4080 TTTTGGCAACAAAGCGCTGCTCTCTGCTGTTTGTGAGGCTCCCTCGAGGCTGGGTGG 4139
Qy 4143 CATCCCTCTCTCTGGGAGTCAAGTCTTCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4202
Db 4140 CATCCCTCTCTCTGGGAGTCAAGTCTTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4199
Qy 4203 GCAGATGCGGATGTTTACAGGAGAGTCAAGGAGCGCTGAGAAACCGAAGAGCAGCAGCC 4262
Db 4200 GCAGATGCGGATGTTTACAGGAGAGTCAAGGAGCGCTGAGAAACCGAAGAGCAGCAGCC 4259
Qy 4263 CCTTCACAGCCTGGGACCTGCATCTGA 4290
Db 4260 CCTTCACAGCCTGGGACCTGCATCTGA 4287

RESULT 4

AAZ52275

ID AAZ52275 standard; cDNA; 4068 BP.

XX

AC AAZ52275;

XX

Qy 3810 CTCTCAGACAGCAATTTTATGGACCTGGTACAGTTCTTCTGACCTTCTTCTCAGTGTGTT 3869
Db 3588 CTCTCAGACAGCAATTTTATGGACCTGGTACAGTTCTTCTGACCTTCTTCTCAGTGTGTT 3647
Qy 3870 CCTCTCTTGTCTCTGCTGGTGTCTGTTTGGAGATCAAAACAAAGTTGTTGGGCTC 3929
Db 3648 CCTCTCTTGTCTCTGCTGGTGTCTGTTTGGAGATCAAAACAAAGTTGTTGGGCTC 3707
Qy 3930 CAGACGTAGAGAGCAATTTCTTCAGAGATGCAACAGATGGCCAGCCGTCCTTGGCTC 3989
Db 3708 CAGACGTAGAGAGCAATTTCTTCAGAGATGCAACAGATGGCCAGCCGTCCTTGGCTC 3767
Qy 3990 TGTAAATGTCCTCTTGGAAACAGATGAGGACCTCTGATCTTATTGGGGGAGTATAAA 4049
Db 3768 TGTAAATGTCCTCTTGGAAACAGATGAGGACCTCTGATCTTATTGGGGGAGTATAAA 3827
Qy 4050 GACTGTTCCTCCAAACCCATTGACATGGAGCCGTGTTTGGCAACAAAGCCGTCGTCCTC 4109
Db 3828 GACTGTTCCTCCAAACCCATTGACATGGAGCCGTGTTTGGCAACAAAGCCGTCGTCCTC 3887
Qy 4110 TGTGTTTGTAGGCTCTCTCAGGCTGGTGGCATCTTCTCAGCAGATGCCGATAGTACAGAGAA 4169
Db 3888 TGTGTTTGTAGGCTCTCTCAGGCTGGTGGCATCTTCTCAGCAGATGCCGATAGTACAGAGAA 3947
Qy 4170 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTACAGAGAGAA 4229
Db 3948 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTACAGAGAGAA 4007
Qy 4230 GTCAGGACCGTGTAGAAACCGGAGAGCAGACGCCCTGTGCACAGCTGGGACCTGCATCTG 4289
Db 4008 GTCAGGACCGTGTAGAAACCGGAGAGCAGACGCCCTGTGCACAGCTGGGACCTGCATCTG 4067
Qy 4290 A 4290
Db 4068 A 4068

RESULT 5
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.
XX
AC AAZ52277;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human soluble attractin-2 cDNA.
XX
KW Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3819
FT /*tag= a
FT /product= "Soluble attractin-2"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA.FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70692.

XX Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 12; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human soluble attractin-2.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease
XX
SQ Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;
Query Match 88.6%; Score 3801.4; DB 3; Length 3819;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3802; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGCTGCGCGCAGCGCGCAACTGAGGCAAGCTGAGGAGGAGGAGCGCGCAGCGCA 60
Db 1 ATGCTGCGCGCAGCGCGCAACTGAGGCAAGCTGAGGAGGAGGAGCGCGCAGCGCA 60
Qy 61 GGCTCTCGCGCGCAGGAGCGCGCGCACTGGGACTGGGAGCTGACAGGCGTGGAGG 120
Db 61 GGCTCTCGCGCGCAGGAGCGCGCGCACTGGGACTGGGAGCTGACAGGCGTGGAGG 120
Qy 121 CCGGGGCTGGGGCGCGGCTGGCGCTCCCGCGCGCTGCTCTCCACCGTGGCGCCACG 180
Db 121 CCGGGGCTGGGGCGCGGCTGGCGCTCCCGCGCGCTGCTCTCCACCGTGGCGCCACG 180
Qy 181 CTGCTGCTGCTGCTGTTGTTGTTCCCGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTGCTGCTGCTGCTGTTGTTGTTCCCGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTG 240
Qy 241 GCGGAGCGCGCGCGCGCGCGCGGCTGCGGCTGCGGCTGCGGCGCGCGCGCGCGCGG 300
Db 241 GCGGAGCGCGCGCGCGCGCGCGGCTGCGGCTGCGGCTGCGGCGCGCGCGCGCGG 300
Qy 301 TGTGACCGCGCGCTGCTCAACGCGCGCTGCTCAACCGCTGGCAACCGCGCGCGCGCTG 360
Db 301 TGTGACCGCGCGCTGCTGCTCAACGCGCGCTGCTCAACCGCTGGCAACCGCGCGCG 360
Qy 361 CCGCGCGCGCTGGGCGGAGCAATGCCAGCACTGGCGGGCGCGCTTCAGACTAACTGGA 420
Db 361 CCGCGCGCGCTGGGCGGAGCAATGCCAGCACTGGCGGGCGCGCTTCAGACTAACTGGA 420
Qy 421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATATAAATACAAAACGAGTGCAGTGG 480
Db 421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATATAAATACAAAACGAGTGCAGTGG 480
Qy 481 CTCATTGAGGACAGGCAAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGAG 540
Db 481 CTCATTGAGGACAGGCAAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGAG 540
Qy 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCT 600
Db 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCT 600
Qy 601 GCATTTAGTGGCTCATTTGTTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660
Db 601 GCATTTAGTGGCTCATTTGTTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660
Qy 661 GCCACATCAGGTTATGCGCTTGTGCAATTTTTTTAGTAGTCTGCTCTATTAATTGACATGGA 720

Db ||||| 661 GCCACATCAGGTTATGCTTGTGCTATTTTATGATGCTGCTTATAATTTGACTGGA 720
QY ||||| 721 TTTAATATTACTTACAGTTTATGATGCTCAATATGCTCAATATGCTCAGGCGGAGGAGTGT 780
Db ||||| 721 TTTAATATTACTTACAGTTTATGATGCTCAATATGCTCAATATGCTCAGGCGGAGGAGTGT 780
QY ||||| 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAATCTGAAAGAGTGAA 840
Db ||||| 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAATCTGAAAGAGTGAA 840
QY ||||| 841 GCATGTGATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db ||||| 841 GCATGTGATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY ||||| 901 TCAAGTGATGTCAGAGGATGCTTCTGCTTCTCAGACTGCGAGGCTCCCTGAGTGTTCAGTT 960
Db ||||| 901 TCAAGTGATGTCAGAGGATGCTTCTGCTTCTCAGACTGCGAGGCTCCCTGAGTGTTCAGTT 960
QY ||||| 961 CTTGTACAGCTTACAGTCAATTTTGGATCTGAGAGGAATATCTTAAGTAAAGCTCCCC 1020
Db ||||| 961 CTTGTACAGCTTACAGTCAATTTTGGATCTGAGAGGAATATCTTAAGTAAAGCTCCCC 1020
QY ||||| 1021 AGAGCATCTCAATAAGCTGCTGCTCAATGGAACCAATATGTCGGTGTGTTGGAGGATATATG 1080
Db ||||| 1021 AGAGCATCTCAATAAGCTGCTGCTCAATGGAACCAATATGTCGGTGTGTTGGAGGATATATG 1080
QY ||||| 1081 TTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTT 1140
Db ||||| 1081 TTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTT 1140
QY ||||| 1141 CCACTTAACCGTCTGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db ||||| 1141 CCACTTAACCGTCTGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY ||||| 1201 AAGGATAAAATTTACATGATGAGGAAATTTGATCAACTGGGAATGACCAATGAG 1260
Db ||||| 1201 AAGGATAAAATTTACATGATGAGGAAATTTGATCAACTGGGAATGACCAATGAG 1260
QY ||||| 1261 TTGAGAGTTTTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db ||||| 1261 TTGAGAGTTTTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY ||||| 1321 CAGTATGAGTGTGAGGCACTCTGACACATGTTTACACTGGAAGATGCGCAGTGGTC 1380
Db ||||| 1321 CAGTATGAGTGTGAGGCACTCTGACACATGTTTACACTGGAAGATGCGCAGTGGTC 1380
QY ||||| 1381 ATGCTGTCATCTTTGGTCACTGGCTCTCTATGATATATAGCAATGTCAGGAATAT 1440
Db ||||| 1381 ATGCTGTCATCTTTGGTCACTGGCTCTCTATGATATATAGCAATGTCAGGAATAT 1440
QY ||||| 1441 GATTTGGATAGACACATGAGTATATACACACCCAGGCTGCTTGTGCGAGGGGT 1500
Db ||||| 1441 GATTTGGATAGACACATGAGTATATACACACCCAGGCTGCTTGTGCGAGGGGT 1500
QY ||||| 1501 TACGGCCATAGCAGTGTGTTACGACCATAGGACCCAGGCTTATACGTTCTATGTTGGCTAC 1560
Db ||||| 1501 TACGGCCATAGCAGTGTGTTACGACCATAGGACCCAGGCTTATACGTTCTATGTTGGCTAC 1560
QY ||||| 1561 AAGGCTTTCAGTGCATTAAGTACCGGCTTGCATGATCTTACCGATATGATGATGAT 1620
Db ||||| 1561 AAGGCTTTCAGTGCATTAAGTACCGGCTTGCATGATCTTACCGATATGATGATGATGAT 1620
QY ||||| 1621 ACCCAGATGTCGACCATTTTAAAGACAGCCGATTTTTCGGTACTTGTGACACAGCTGTG 1680
Db ||||| 1621 ACCCAGATGTCGACCATTTTAAAGACAGCCGATTTTTCGGTACTTGTGACACAGCTGTG 1680
QY ||||| 1681 ATAGTGTGGAACCATGCTGGTGTGTTGGGGGAAACACACATGATGATGATGATGATGATGAT 1740
Db ||||| 1681 ATAGTGTGGAACCATGCTGGTGTGTTGGGGGAAACACACATGATGATGATGATGATGATGAT 1740
QY ||||| 1741 CATGGCCCAATGCTTCTCTCAGATTTTCAATGCTTATGATGATGATGATGATGATGATGATGAT 1800

Db ||||| 1741 CATGGCCCAATGCTTCTCTCAGATTTTCAATGCTTATGATGATGATGATGATGATGATGATGAT 1800
QY ||||| 1801 TCACTGCTTCCAGAGCTGATCTCCACCATGATGATCAACAGATTTGGCCATTCAGCAGTC 1860
Db ||||| 1801 TCACTGCTTCCAGAGCTGATCTCCACCATGATGATCAACAGATTTGGCCATTCAGCAGTC 1860
QY ||||| 1861 TTAACAACAAGCAGCAGTATGATGTTTGGTGGTTCCTCAATGATCTCTCTCTCAGCAGATC 1920
Db ||||| 1861 TTAACAACAAGCAGCAGTATGATGTTTGGTGGTTCCTCAATGATCTCTCTCTCAGCAGATC 1920
QY ||||| 1921 CTGGTATTCACCTCGGAAACAGTGTGATGCGATCGAGTGAAGCCGCTTGTGTACAGCA 1980
Db ||||| 1921 CTGGTATTCACCTCGGAAACAGTGTGATGCGATCGAGTGAAGCCGCTTGTGTACAGCA 1980
QY ||||| 1981 GGAACCTGGTATTCGGTGTGTTGGAAACAGAGGTCTCTCAGTGTATCTCTGGGCGGTG 2040
Db ||||| 1981 GGAACCTGGTATTCGGTGTGTTGGAAACAGAGGTCTCTCAGTGTATCTCTGGGCGGTG 2040
QY ||||| 2041 GCACCTGATGAACAGAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTTGAC 2100
Db ||||| 2041 GCACCTGATGAACAGAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTTGAC 2100
QY ||||| 2101 CATGACAGATGTGACAGCAGACACAGATTTTACAGCTGTACAGCCAAACCAATGACTGC 2160
Db ||||| 2101 CATGACAGATGTGACAGCAGACACAGATTTTACAGCTGTACAGCCAAACCAATGACTGC 2160
QY ||||| 2161 CACTGCTGCAATGACCAATTTGCTCCAGGAAACACAGCTGTCCAGAGGCGCAGATCTCC 2220
Db ||||| 2161 CACTGCTGCAATGACCAATTTGCTCCAGGAAACACAGCTGTCCAGAGGCGCAGATCTCC 2220
QY ||||| 2221 ATTTTGTAGTATGAGAAATTTGCTCCAGGAAACACAGCTGTCCAGAGGCGCAGATCTCC 2280
Db ||||| 2221 ATTTTGTAGTATGAGAAATTTGCTCCAGGAAACACAGCTGTCCAGAGGCGCAGATCTCC 2280
QY ||||| 2281 AGCTGACAGAGTGTGCTCCAGCAGAACTGCTCCAGGAGGCGCAGGAAATCAGAGTGC 2340
Db ||||| 2281 AGCTGACAGAGTGTGCTCCAGCAGAACTGCTCCAGGAGGCGCAGGAAATCAGAGTGC 2340
QY ||||| 2341 ATGCTGCTCCGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db ||||| 2341 ATGCTGCTCCGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
QY ||||| 2401 AAAATTTACTTACTGCTCCAGGAGAAATTTATGACAAATGCTTAAATTTGCTGAGAAACCAAT 2460
Db ||||| 2401 AAAATTTACTTACTGCTCCAGGAGAAATTTATGACAAATGCTTAAATTTGCTGAGAAACCAAT 2460
QY ||||| 2461 GCTCTTTTGGTCTCTTCAACCCAGAGAGTGTGCTTCTTCTTAAAGCAGCTGCGA 2520
Db ||||| 2461 GCTCTTTTGGTCTCTTCAACCCAGAGAGTGTGCTTCTTCTTAAAGCAGCTGCGA 2520
QY ||||| 2521 ATAAATGCTGCTCTCAGAGCAGTGTCCAGCTTAAACCCAGTGTGCTGCTGCTGCTGCTGCT 2580
Db ||||| 2521 ATAAATGCTGCTCTCAGAGCAGTGTCCAGCTTAAACCCAGTGTGCTGCTGCTGCTGCTGCT 2580
QY ||||| 2581 AAGATCAATGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Db ||||| 2581 AAGATCAATGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
QY ||||| 2641 CAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db ||||| 2641 CAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
QY ||||| 2701 ACTGGGAGCTGAAGCTGCAACCTGCTATCAACCCACTCAATGGTGTGCTGCTGCTGCTGCT 2760
Db ||||| 2701 ACTGGGAGCTGAAGCTGCAACCTGCTATCAACCCACTCAATGGTGTGCTGCTGCTGCTGCT 2760
QY ||||| 2761 CTTGCAAAACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db ||||| 2761 CTTGCAAAACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
QY ||||| 2821 GATTCGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Db ||||| 2821 GATTCGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880

QY 243 CGAGCGCGCGCGCGCGCGTCTCGGGCTCAGCGCAGCGCGAGGCCAGGAATG 302
Db 240 CGAGCGCGCGCGCGCGCGCGTCTCGGGCTCAGCGCAGCGCGAGGCCAGGAATG 299
QY 303 TGACCGCGCGCGTGTCTCAACGGCGCGTGTCTGCAACCGCTGGCACCGCGCAGTGTGCGCC 362
Db 300 TGACCGCGCGCGTGTCTCAACGGCGCGTGTCTGCAACCGCTGGCACCGCGCAGTGTGCGCC 359
QY 363 CGCGCGCGTGTGTGCGCGAGCAATGCCAGCACTGCGGGCGCGCTTACAGCTAACTGATC 422
Db 360 CGCGCGCGTGTGTGCGCGAGCAATGCCAGCACTGCGGGCGCGCTTACAGCTAACTGATC 419
QY 423 TTCTCGGGTTTGTGACAGATGACCTGGAAATTTAAATACAAAACGAAGTGCACGTGGCT 482
Db 420 TTCTCGGGTTTGTGACAGATGACCTGGAAATTTAAATACAAAACGAAGTGCACGTGGCT 479
QY 483 CATTTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATTTTGTCTACAGATG 542
Db 480 CATTTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATTTTGTCTACAGATG 539
QY 543 TAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGTGTC 602
Db 540 TAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGTGTC 599
QY 603 ATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGTGTGTC 662
Db 600 ATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGTGTGTC 659
QY 663 CACATCAGTTATGCTTGTGCTGCTATTTTATGATGCTGCTTATTAATTTGCTGAT 722
Db 660 CACATCAGTTATGCTTGTGCTGCTATTTTATGATGCTGCTTATTAATTTGCTGAT 719
QY 723 TAATTAATTTACAGTTTGTATGCTGCAAAATPAACTGCTCAGCGCGAGAGAGTGTAA 782
Db 720 TAATTAATTTACAGTTTGTATGCTGCAAAATPAACTGCTCAGCGCGAGAGAGTGTAA 779
QY 783 GATCAGTAATAGCAGCAACTGTTGAAATGTGAATGTTCTGAAACTGGAAGAGTGAAGC 842
Db 780 GATCAGTAATAGCAGCAACTGTTGAAATGTGAATGTTCTGAAACTGGAAGAGTGAAGC 839
QY 843 ATGTGACATTCCTCAGTGTACAGCAACTGTGGTTTCTCCTCAGCGGCACTGCAATTC 902
Db 840 ATGTGACATTCCTCAGTGTACAGCAACTGTGGTTTCTCCTCAGCGGCACTGCAATTC 899
QY 903 AAGTGTGTGAGAGATGCTTCTGCTTCTCAGACTGGCAGGGCTCGGATGTTCAGTTC 962
Db 900 AAGTGTGTGAGAGATGCTTCTGCTTCTCAGACTGGCAGGGCTCGGATGTTCAGTTC 959
QY 963 TGTACCACTAACCAAGTCAATTTGGACTTCGAGAGGAATTTCTAACTTAAAGCTCCCGAG 1022
Db 960 TGTACCAAGCTAACCAAGTCAATTTGGACTTCGAGAGGAATTTCTAACTTAAAGCTCCCGAG 1019
QY 1023 AGCATCTCATAAGCTGTGTCAATGGAACAATATGTTGGGTTTGTGGAGATATATGTT 1082
Db 1020 AGCATCTCATAAGCTGTGTCAATGGAACAATATGTTGGGTTTGTGGAGATATATGTT 1079
QY 1083 CAACCACTCAGATTAATCAATGTTTGTAGCGTATGACCTTCTTACGAGTGGCTTCC 1142
Db 1080 CAACCACTCAGATTAATCAATGTTTGTAGCGTATGACCTTCTTACGAGTGGCTTCC 1139
QY 1143 ACTAAACCGTTCTGTGAACAATGTTGTTGTAGATATGTTCTTCTTGGCATTAACAA 1202
Db 1140 ACTAAACCGTTCTGTGAACAATGTTGTTGTAGATATGTTCTTCTTGGCATTAACAA 1199
QY 1203 GGATAAATTTATCATGTATGGAGGAAATTTGATTCNACTGGGAATGTGACCAATGATTT 1262
Db 1200 GGATAAATTTATCATGTATGGAGGAAATTTGATTCNACTGGGAATGTGACCAATGATTT 1259
QY 1263 GAGAGTTTTTCACTTCAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
Db 1260 GAGAGTTTTTCACTTCAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
QY 1323 GTATGCACTGCTGGGCACTCTGCAACATGTTGTTACACTGAGAAATGGCGGAGTGTCTAT 1382

Db 1320 GTATGCACTGCTGGGCACTCTGCAACATGTTGTTACACTGAGAAATGGCGGAGTGTCTAT 1379
QY 1383 GCTGTGTCATTTTGTGTCTGCTGCTCTCTATGGAATATAAGCAATGTGCAAGGAATATGA 1442
Db 1380 GCTGTGTCATTTTGTGTCTGCTGCTCTCTATGGAATATAAGCAATGTGCAAGGAATATGA 1439
QY 1443 TTTTGATATAGAACACATGAGTGTATATTACACCCAGGGTCCCTTGTGCAAGGGGTTA 1502
Db 1440 TTTTGATATAGAACACATGAGTGTATATTACACCCAGGGTCCCTTGTGCAAGGGGTTA 1499
QY 1503 CGGCCATAGCAGTGTTTACGACCATAGACCCAGGCCCCCTATACGTTTCAATGTTGCTACAA 1562
Db 1500 CGGCCATAGCAGTGTTTACGACCATAGACCCAGGCCCCCTATACGTTTCAATGTTGCTACAA 1559
QY 1563 GCTTTTCAGTGCCATATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATAC 1622
Db 1560 GCTTTTCAGTGCCATATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATAC 1619
QY 1623 CCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGAT 1682
Db 1620 CCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGAT 1679
QY 1683 AGTGTGTGGAACCATGCTGTTTGGGGGAAAACACACAAATGACACATCTATGAGCCA 1742
Db 1680 AGTGTGTGGAACCATGCTGTTTGGGGGAAAACACACAAATGACACATCTATGAGCCA 1739
QY 1743 TGGCGCCAAATGCTTCTCTCAGATTTTCATGCGCTATGACATTCCTGTGACCGTGTCTC 1802
Db 1740 TGGCGCCAAATGCTTCTCTCAGATTTTCATGCGCTATGACATTCCTGTGACCGTGTCTC 1799
QY 1803 AGTGTTCCTCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862
Db 1800 AGTGTTCCTCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1859
QY 1863 ACACAAACAGCAGCATGATGTTTGGTGGTTCATAGTCTCTCCTCAGCAGCATCTCT 1922
Db 1860 ACACAAACAGCAGCATGATGTTTGGTGGTTCATAGTCTCTCCTCAGCAGCATCTCT 1919
QY 1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTGTGTAGCAGCAGG 1982
Db 1920 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTGTGTAGCAGCAGG 1979
QY 1983 ACCTGGTATTCGGTGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCTGGGCGCTGGC 2042
Db 1980 ACCTGGTATTCGGTGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCTGGGCGCTGGC 2039
QY 2043 AACTGATGAAACAAAGAAAGAAATGTTAAATCAGAAATGTTTTTCCAAAGAACTTTGACCA 2102
Db 2040 AACTGATGAAACAAAGAAAGAAATGTTAAATCAGAAATGTTTTTCCAAAGAACTTTGACCA 2099
QY 2103 TGACAGATGTGACACAGCAACAGATTTGTTACAGTGTATGACGCCAACCAATGATGCTGCA 2162
Db 2100 TGACAGATGTGACACAGCAACAGATTTGTTACAGTGTGACGCCAACCAATGATGCTGCA 2159
QY 2163 CTGGTGAATGACCATTTGTGTCGCCAGGAACACAGCTGTCTCAGAGGCCAGATCTCCAT 2222
Db 2160 CTGGTGAATGACCATTTGTGTCGCCAGGAACACAGCTGTCTCAGAGGCCAGATCTCCAT 2219
QY 2223 TTTTAGTGTATGAGAAATGCCCCAAGGATAACCCCATGTACTGTATCTGTAAACAAAGAACAG 2282
Db 2220 TTTTAGTGTATGAGAAATGCCCCAAGGATAACCCCATGTACTGTATCTGTAAACAAAGAACAG 2279
QY 2283 CTGACAGGCTGTGCGCTGACAGCAAGTGTGCGGAGCCCGGAAATCAGGAGTGCAT 2342
Db 2280 CTGACAGGCTGTGCGCTGACAGCAAGTGTGCGGAGCCCGGAAATCAGGAGTGCAT 2339
QY 2343 TGCCCTGCCCCGAAATATCTGTGCGCATTTGGTGGAACTCATGTTTGA 2402
Db 2340 TGCCCTGCCCCGAAATATCTGTGCGCATTTGGTGGAACTCATGTTTGA 2399
QY 2403 AATTACTACTGCCAAGGAGAAATATGACATGTCTAAATTTGTTCTGTAGGACCAACATGC 2462

1872	Db		AGTTCTCCCGAGACCTGAGCTCCATCAGATGTCACCGAGTTTGGCCACTCAGCAGTCTT	1931
1863	Qy		ACAACAGCACCATGTATGTTTCGGTGGTTTCAATAGTCTCCTCTCAGCGACATCCT	1922
1932	Db		GCACAAGCACCATGTATGTTTCGGCGGCTTCAACAGCCTCCTCTCAGTGAAGTCTT	1991
1923	Qy		GGTATTCACTCCGGAACAGTGTGAGTGGGCAATCGGAGTGAAGCGCGCTGTTTACGACGAG	1982
1992	Db		AGTCTTCACTCCGAGCAGTGTGATGCACCCGAGCGAAGCTGCTTGTGTGCACGACGAG	2051
1983	Qy		ACCTGGTATTTCGGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGC	2042
2052	Db		ACCTGGCATCCGGTGTCTGTGGGACACACAGTGGTCTCGATGTACCTCTCGGAGTTGGC	2111
2043	Qy		AACGTATGAACAAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCA	2102
2112	Db		AACGTGAAGAACAGACAGAAAAGTTAAATCAGAAATGTTTCTCTAAAAGAACCTTGACCA	2171
2103	Qy		TGACAGATGTGACACAGCACACAGATTGTATACGCTGTACAGCCACCAACCAATGACCTGCCA	2162
2172	Db		TGACAGATGTGACACAGCACACAGATTGTATACGCTGTACAGCCCAATACCAATGACCTGCCA	2231
2163	Qy		CTGGTGCAATGACCAATTTGTGTCCCCAGAAACACAGCTGCTCAGAAGCCAGATCTCCAT	2222
2232	Db		CTGGTGCAATGATCACTGTGTCTCTGTGAACACAGCTGCACAGAAGCCAGATCTCCAT	2291
2223	Qy		TTTTAGTATGAGAAATTTGCCCAAGGATACCCCATGTACTACTGTATCAAGAGACCCAG	2282
2292	Db		TGCCAAGTATGACAAATTTGCCCCAGGATACCCCATGTACTACTGTCAATGAAGAAACCCAG	2351
2283	Qy		CTGCAGAGCTGTGCCCTGGACACAGAACTGCAGTGGGAGCCCCGGAAATCAGAGATGTCAT	2342
2352	Db		CTGCAGAGCTGTGCCCTAGATCAGATTGCCAGTGGAGACCTCGAAATCAAGAGTGCAT	2411
2343	Qy		TGCCCTGCCGAAAATATCTGTGGCATTTGGCTGGCAATTTGGTGTGAAAATCATGTTTGAA	2402
2412	Db		CGCCCTGCCGAAAACATCTGTGGCATTTGGCTGGCAATTTGGCTGGAAATCTGTGTCTGAA	2471
2403	Qy		AATTTACTTGCCTAAGAGAAATATGACAAATGCTTAAATTTGCTGTAGGAACACCAATGC	2462
2472	Db		AATCACTACTGCTAAGAGAAATATGACAAATGCTTAAATTTGCTGTAGGAACACCAATGC	2531
2463	Qy		CCTTTTCGGCTTCTTTACAAACCCAGAGAAGGTAGAAATTTGCTCTTAAGCAGCTGGAAAT	2522
2532	Db		CTTTTTCGGCTTCGGCTTACGTCCCAAGAGAAGTGGAGTTTGTCTTAAGCAGCTTCGATT	2591
2523	Qy		AATCGAGTCACTCAGAGCATGTCACAGCTCACCTTAACCCCATGGCTGGCCCTCGGAA	2582
2592	Db		AATGCAATCATCTCAGAGCAGTCCAGCTCACTCTGACTCCATGGGTTGGTCTTTCGGAA	2651
2583	Qy		GATCAATGTGTCTTACTGGTCTGGGAGATATGCCCAATTTACAATATGTTTACTACA	2642
2652	Db		GATCAATGTATCTTACTGGTCTGGGAGATATGCTCCAATTCACAAATAGTTTGCTGCA	2711
2643	Qy		GTGATGCGGTCTGAGCCCAAGTATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTAC	2702
2712	Db		GTGATGCCATCTGAGCCCAAGTATGCGGCTTCTGTGGGATCTTGTGCAGAGCTTAGTAC	2771
2703	Qy		TCGGGACTGAAGGCTCAACCTGCATCAACCCCACTCAATGGTATGCTGTGAAAAGGCC	2762
2772	Db		TCGGGACTGAAGGCTCAACCTGCATCAACCCCACTCAATGGAGTGTTTGTGAAAAGGCC	2831
2763	Qy		TGCAAAACACAGTCTAAGCAGTCCCGGACACCATGTCCTTGAAGACAGCATGTGAGAA	2822
2832	Db		TGCAAAACACAGTCCCAAGCAGTCCCGGACACCCCTGTGCCCTGCGGACAGCATGTGTGTA	2891
2823	Qy		TTGCACAGCGGACGCTCTGAGTGCATGTGTGTGACAGCAACATGAAGCAGTGTGTGGAATC	2882
2892	Db		GTGCACAGCGAGAGCTCCGAATGCATGTGTGTGACAGCAACATGAAGCAGTGTGTGGAATC	2951
2883	Qy		CAATGCCCTATGTGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGACAC	2942

QY 4023 TCCTGATCTTATTTGGGGAGTATAAGACTGTTCCTCCAAACCCATTTGACTGGAGCCGTG 4082
 |||||
 Db 4092 TCCTGATCTCATTTGGGGAGTATAAGACTGTTCCTTAAGCCCATTTGCCCTGGAGCCCTG 4151
 |||||
 QY 4083 TTTTGGCAACAAAGCGCTGCTCTCTGTTGTTGTGAGGCTCCCTCGAGCCCTGGTGG 4142
 |||||
 Db 4152 CTTTGGTAAACAGCTGCTGCTCTCTGTTGTTGTGAGGCTCCCTCGAGGACTGGGAGG 4211
 |||||
 QY 4143 CATCCCTCTCTCGGCGAGTCAGGTCCTTGTGTCGTCGCGCCAGCCCTGGTGGACATTTCTCA 4202
 |||||
 Db 4212 AATCCCTCTCTCGTGTGCTGAGTCAGGTCCTTGTGTCGTCGCGCCAGTGCCTGGTGGACATTTCTCA 4271
 |||||
 QY 4203 GCAGATGCGGTAGTGTACAGGAGAGTCAAGGAGCCCTGAGAACCGGAGGAGGAGCC 4262
 |||||
 Db 4272 GCAGATGCCAATAGTGTACAGGAGAGTCAAGGAGCTGTAGGAACCGGAGGAGGAGCC 4331
 |||||
 QY 4263 CCCTGCACAGCCTGGGACCTGCATCTGA 4290
 |||||
 Db 4332 CCCTGCACAGCCTGGAACCTGCATTTGA 4359
 |||||

RESULT 8

AAZ52274
 ID AAZ52274 standard; cDNA; 3597 BP.

AC AAZ52274;

DT 18-JUL-2000 (first entry)

XX Human soluble attractin-1 cDNA.

XX Human; soluble attractin-1; immune response; macrophage; monocyte;
 KW T cell; immunostimulant; immunosuppressed patient; cancer; ss.
 KW immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT 1. 3597

FT /tag= a

FT /product= "Soluble attractin-1"

XX WO200015651-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-US020948.

XX 14-SEP-1998; 98US-0100137P.

XX (DAND) DANA FARBER CANCER INST INC.

XX Duke-Cohan JS, Schlosseman SF;

XX WPI; 2000-271373/23.

XX P-PSDB; AAY70689.

XX Isolated nucleic acids encoding human attractin polypeptides useful for
 enhancing immune responses.

XX Claim 3; Fig 8; 120pp; English.

XX The patent discloses four forms of human attractin polypeptides which
 enhance immune response by promoting macrophage and monocyte spreading in
 the presence of T cells. These include soluble attractin-1 and -2 and
 membrane attractin-1 and -2. These various forms of attractin are encoded
 by alternatively spliced mRNA molecule transcribed from a single gene.
 CC The present sequence is a cDNA encoding soluble attractin-1 obtained from
 activated T cell and human fetal liver libraries. This sequence can be
 used to enhance immune response in immunosuppressed patients such as
 those undergoing chemo- and radio-therapy treatment for cancer or those
 suffering from common variable immunodeficiency syndrome. The proteins

CC may also be used to screen modulators (agonists and antagonists) of
 CC immune responses which may also be used to regulate immune reactions.
 CC Attractin antibodies can be used to inhibit immune response in transplant
 CC recipients or patients afflicted with autoimmune disease

XX Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 80.9%; Score 3469.4; DB 3; Length 3597;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 3507; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

QY 270 GTGGGCTCAGCGGAGCGAGGCAAGAAATGTACCGGCCCTGTGTCAACGCGGTGCG 329
 |||||
 Db 48 GGGGGGAGCGGAGCGCTCGCGGGCAGGAGCGGCGCGCACTGTGTCAACGCGGTGCG 107
 |||||
 QY 330 CTGCAACCTTGGCAACCGGCACTGTGCTGCTCCCGCGGCTGGTGGGCGAGCAATGGCA 389
 |||||
 Db 108 CTGCAACCTTGGCAACCGGCACTGTGCTGCTCCCGCGGCTGGTGGGCGAGCAATGGCA 167
 |||||
 QY 390 GCACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTTGTGAGAGATGGACCTGG 449
 |||||
 Db 168 GCACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTTGTGAGAGATGGACCTGG 227
 |||||
 QY 450 AAATTATAAATAACAAACGAAGTGCACCTGGCTCTATTGAAGGACAGCCAAATAGATAAT 509
 |||||
 Db 228 AAATTATAAATAACAAACGAAGTGCACCTGGCTCTATTGAAGGACAGCCAAATAGATAAT 287
 |||||
 QY 510 GAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGACCAATTTATATTATGA 569
 |||||
 Db 288 GAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGACCAATTTATATTATGA 347
 |||||
 QY 570 TGGGGACTCAATTTATGACCGCTAGTTGCTGCTCAATTTAGTGGCCTCAATTTCTCGAGAG 629
 |||||
 Db 348 TGGGGACTCAATTTATGACCGCTAGTTGCTGCTCAATTTAGTGGCCTCAATTTCTCGAGAG 407
 |||||
 QY 630 AGATGGCAATGAGACTGTGCTGAGGTTGTGCGCACATCAGAGTTATGCTGCTGCAATTT 689
 |||||
 Db 408 AGATGGCAATGAGACTGTGCTGAGGTTGTGCGCACATCAGAGTTATGCTGCTGCAATTT 467
 |||||
 QY 690 TTTTGTAGTGTGCTGCTTATTAATTTGACTGGATTTAATTAATTTACTACAGTTTGTATGTG 749
 |||||
 Db 468 TTTTGTAGTGTGCTGCTTATTAATTTGACTGGATTTAATTAATTTACTACAGTTTGTATGTG 527
 |||||
 QY 750 TCCAAATAACTGCTCAGGCCGAGGAGTGTAAAGATCAGTAATAGCAGCAAACTGTTGA 809
 |||||
 Db 528 TCCAAATAACTGCTCAGGCCGAGGAGTGTAAAGATCAGTAATAGCAGCAAACTGTTGA 587
 |||||
 QY 810 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTTCTCACTGTACAGACAA 869
 |||||
 Db 588 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTTCTCACTGTACAGACAA 647
 |||||
 QY 870 CTGTGGTTTTCTCATCGAGGCACTGCAATTTCAAGTGTATGTCAGAGGATGCTCTCTGCTT 929
 |||||
 Db 648 CTGTGGTTTTCTCATCGAGGCACTGCAATTTCAAGTGTATGTCAGAGGATGCTCTCTGCTT 707
 |||||
 QY 930 CTCAGACTGGCAGGCTCTGATGTTTCAAGTGTATGTCAGGTTACCACTACCACTATTTGGAC 989
 |||||
 Db 708 CTCAGACTGGCAGGCTCTGATGTTTCAAGTGTATGTCAGGTTACCACTACCACTATTTGGAC 767
 |||||
 QY 990 TCGAGAGCAATATTTCAACTTAAAGCTCCCGAGGCACTCTCATAAAGCTGTGGTCAATGG 1049
 |||||
 Db 768 TCGAGAGCAATATTTCAACTTAAAGCTCCCGAGGCACTCTCATAAAGCTGTGGTCAATGG 827
 |||||
 QY 1050 AAACATTTATGGGGTGTGGAGGATATATGTTCAACCACTCAGATTAATACATGGTTCT 1109
 |||||
 Db 828 AAACATTTATGGGGTGTGGAGGATATATGTTCAACCACTCAGATTAATACATGGTTCT 887
 |||||
 QY 1110 ACGGTATGACCTTGTCTTCTAGGGAGTGGCTTCCCACTAAACCGTTCTGTGAACAATGGT 1169
 |||||
 Db 888 ACGGTATGACCTTGTCTTCTAGGGAGTGGCTTCCCACTAAACCGTTCTGTGAACAATGGT 947
 |||||
 QY 1170 TGTATGATATGCTCATTTTGGCATTATACAAGATAAAATTTACATGTTATGAGGAAA 1229
 |||||

QY 3389 AGGGCTCAAGGGGACGAGTCCAGCTATCTGAGGTAGAAATCGATACCAAGAAACC 3448
DB |||||||
QY 3167 AGGGCGTCAAGGGGACGAGTCCAGCTATCTGAGGTAGAAATCGATACCAAGAAACC 3226
DB |||||||
QY 3449 CTCTCAGAGGACATGTTATTATATCTCTTCTTATTGACTATCATCTTACCTTTAGTCTAT 3508
DB |||||||
QY 3227 CTCTCAGAGGACATGTTATTATATCTCTTCTTATTGACTATCATCTTACCTTTAGTCTAT 3286
DB |||||||
QY 3509 CCCAGGAAGATGCTGCTATTACACAGCTATCAATTTTGGGCTACTCTCTGACGACAAA 3568
DB |||||||
QY 3287 CCCAGGAAGATGCTGCTATTACACAGCTATCAATTTTGGGCTACTCTCTGACGACAAA 3346
DB |||||||
QY 3569 ACAGGGATTGGACATGTTCAATCAATGCTCCCAAGAAATTTCAACCTCAACATCACCTGGG 3628
DB |||||||
QY 3347 ACAGGGATTGGACATGTTCAATCAATGCTCCCAAGAAATTTCAACCTCAACATCACCTGGG 3406
DB |||||||
QY 3629 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTGTTTTCATAAACC 3688
DB |||||||
QY 3407 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTGTTTTCATAAACC 3466
DB |||||||
QY 3689 ACATTAAAGGCTACAAAGATAGTTTCTCTAATGAGAGAGTTTGATTTTCGCAACCCCAA 3748
DB |||||||
QY 3467 ACATTAAAGGCTACAAAGATAGTTTCTCTAATGAGAGAGTTTGATTTTCGCAACCCCAA 3526
DB |||||||
QY 3749 ATATCACTTTCTTTGTTTATGCTAGTAATTTTCACTGGCCCATCAAAATTCAGAT 3803
DB |||||||
QY 3527 ATATCACTTTCTTTGTTTATGCTAGTAATTTTCACTGGCCCATCAAAATTCAGAT 3581
DB |||||||

RESULT 9

AA572657
ID AAS72657 standard; cDNA; 3597 BP.
AC AA572657;
XX
DT 13-FEB-2002 (first entry)
DE
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG08470.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 8461; 103pp; English.
XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 80.9%; Score 3469.4; DB 5; Length 3597;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 3507; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

QY 270 GTGCGGCTCAGCGGACGCGGAGGCAAGAAATGTGACCGGCCCTGTGTCAACGGGCTCG 329

DB 48 GCGCGGCGACGCGGAGCGCTCGCGGGCAGGAGCGGGCGGCACCTGTGTCAACGGGCTCG 107

QY 330 CTGCAACCTGSCACCGCCAGTGGCTCTGCCCGCGGCTGGGTGGCGGAGCAATGCCA 389

DB 108 CTGCAACCTGSCACCGCCAGTGGCTCTGCCCGCGGCTGGGTGGCGGAGCAATGCCA 167

QY 390 GCATCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGACCTGG 449

DB 168 GCATCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGACCTGG 227

QY 450 AAATTATATAAATACAAACGAAGTCAGCTGCTCATTTGAAGGACAGCCAAATAGAAAT 509

DB 228 AAATTATATAAATACAAACGAAGTCAGCTGCTCATTTGAAGGACAGCCAAATAGAAAT 287

QY 510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGA 569

DB 288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGA 347

QY 570 TGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCCCTCATTTGCTCTGAGAG 629

DB 348 TGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCCCTCATTTGCTCTGAGAG 407

QY 630 AGATGGCAATGAGACTGTCCCTGAGGTTTGTGCCACATCAGGTTATGCTTGTGCTGCAATTT 689

DB 408 AGATGGCAATGAGACTGTCCCTGAGGTTTGTGCCACATCAGGTTATGCTTGTGCTGCAATTT 467

QY 690 TTTTAGTCAATCTGCTTATAATTTTACTGGAATTTTAAATTTACTTTACAGTTTGTGATGTG 749

DB 468 TTTTAGTCAATCTGCTTATAATTTTACTGGAATTTTAAATTTACTTTACAGTTTGTGATGTG 527

QY 750 TCCAAATTAATCTGCTCAGGCGGAGGAGTGAAGTCAATAGTAAAGGAGGAACTGTGTA 809

DB 528 TCCAAATTAATCTGCTCAGGCGGAGGAGTGAAGTCAATAGTAAAGGAGGAACTGTGTA 587

QY 810 ATGTGAATGTTTCTGAAAACCTGAAAAGGTGAAGCATGTGACATTTCTTCACTGTACAGACAA 869

DB 588 ATGTGAATGTTTCTGAAAACCTGAAAAGGTGAAGCATGTGACATTTCTTCACTGTACAGACAA 647

QY 870 CTGTGGTTTCTCTCATCGAGGCATCTGCAATTTCAAGTGAATGTGACAGATGCTCTCTT 929

DB 648 CTGTGGTTTCTCTCATCGAGGCATCTGCAATTTCAAGTGAATGTGACAGATGCTCTCTT 707

QY 930 CTCAGACTGGCAGGCTCTGATGTTTCACTTCTCTGTACACAGCTAACCCAGTCATTTTGGAC 989

DB 708 CTCAGACTGGCAGGCTCTGATGTTTCACTTCTCTGTACACAGCTAACCCAGTCATTTTGGAC 767

QY 990 TCGAGAGGAATATTCTTAACCTTAAAGCTCCCGACAGCATCTCATAAAGCTGTGTCATGG 1049

DB 768 TCGAGAGGAATATTCTTAACCTTAAAGCTCCCGACAGCATCTCATAAAGCTGTGTCATGG 827

QY 1050 AAACATTTATGCGGTTGTTGGAGATATATGTTCAACCACTCAGATTATAAACAATGTTCT 1109
DB 828 AAACATTTATGCGGTTGTTGGAGATATATGTTCAACCACTCAGATTATAAACAATGTTCT 887
QY 1110 AGCGATATGACCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGGAACAATGTGGT 1169
DB 888 AGCGATATGACCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGGAACAATGTGGT 947
QY 1170 TGTATGATATGTTGATTTGTCATTTATCAAGGATATAAATTTACATGTATGAGGAAA 1229
DB 948 TGTATGATATGTTGATTTGTCATTTATCAAGGATATAAATTTACATGTATGAGGAAA 1007
QY 1230 AATTGATCCAACTGGGAATGTGACCAATGATTTGAGAGTTTTCACATTCATAATGAGTC 1289
DB 1008 AATTGATTTCACTGGGAATGTGACCAATGATTTGAGAGTTTTCACATTCATAATGAGTC 1067
QY 1290 ATGGGTGTTGTTGACCCCTTAAGCAAGAGAGAGTATGCACTGGTGGGCACTCTGCACA 1349
DB 1068 ATGGGTGTTGTTGACCCCTTAAGCAAGAGAGAGTATGCACTGGTGGGCACTCTGCACA 1127
QY 1350 CATTTGTTACACTGAAGAAATGGCCGAGTGGTCATGCTGGTCACTTCTGGTCACTGGCCCTCT 1409
DB 1128 CATTTGTTACACTGAAGAAATGGCCGAGTGGTCATGCTGGTCACTTCTGGTCACTGGCCCTCT 1187
QY 1410 CTATGATATATAAGCAATGTGACAGGAATATGATTTGGATAGAAACAACATGGAGTATAT 1469
DB 1188 CTATGATATATAAGCAATGTGACAGGAATATGATTTGGATAGAAACAACATGGAGTATAT 1247
QY 1470 ACACACCCAGGTCGCCCTTGTCAAGGGGTTACGGCCATAGCAGATGTTTACCAACCATAG 1529
DB 1248 ACACACCCAGGTCGCCCTTGTCAAGGGGTTACGGCCATAGCAGATGTTTACCAACCATAG 1307
QY 1530 GACCAGGCGCCCTATACGTTTATGTTGGCTACAAGGCTTTTCACTGCCAATAGTACCGGCT 1589
DB 1308 GACCAGGCGCCCTATACGTTTATGTTGGCTACAAGGCTTTTCACTGCCAATAGTACCGGCT 1367
QY 1590 TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGACCAATCTTAAAGACAG 1649
DB 1368 TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGACCAATCTTAAAGACAG 1427
QY 1650 CCGATTTTTCGGTTTACCTTGACACAGAGCTGTGATAGTGGAGTGGACCACTGCTGGTGTGG 1709
DB 1428 CCGATTTTTCGGTTTACCTTGACACAGAGCTGTGATAGTGGAGTGGACCACTGCTGGTGTGG 1487
QY 1710 GGGAAAACACACAATGACACATCTATGAGCCATGGCGCCCAATGTTCTTTCAGATTT 1769
DB 1488 GGGAAAACACACAATGACACATCTATGAGCCATGGCGCCCAATGTTCTTTCAGATTT 1547
QY 1770 CATGSCCTATGACATTTGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA 1829
DB 1548 CATGSCCTATGACATTTGCTGTGACCGCTGTGATGCTTCCAGACCTGAT-TCCACCA 1606
QY 1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTTCACACAACAGACCACTATGTGTGTG 1888
DB 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTTCACACAACAGACCACTATGTGTGTG 1666
QY 1889 GTGGTTTCAATAGTCTCTCTCAGGACATCTCTGTTATTCACCTCGGAACAGTGTATG 1948
DB 1667 GTGGTTTCAATAGTCTCTCTCAGGACATCTCTGTTATTCACCTCGGAACAGTGTATG 1726
QY 1949 CGCATCGGAGTGAAGCCGCTTGTGATGACAGACCTGTTATTCGGTGTGTGTGGAACA 2008
DB 1727 CGCATCGGAGTGAAGCCGCTTGTGATGACAGACCTGTTATTCGGTGTGTGTGGAACA 1786
QY 2009 CAGGTCGTCTCAGTGTATCTCGTGGGCTGGCACTGATGAACAAGAGAAAGTTAA 2068
DB 1787 CAGGTCGTCTCAGTGTATCTCGTGGGCTGGCACTGATGAACAAGAGAAAGTTAA 1846
QY 2069 AATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATT 2128
DB 1847 AATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATT 1906

QY 2129 GTTACAGCTGTACAGCCAAACCAATGATGCTGCCACTGGTGCATTAAGCAATGTGTGCCCA 2188
DB 1907 GTTACAGCTGTACAGCCAAACCAATGATGCTGCCACTGGTGCATTAAGCAATGTGTGCCCA 1966
QY 2189 GGAACCAACAGCTGTCTCAGAGGCCAGATCTCCATTTTATGATGTAGAAATGCCCCAAGG 2248
DB 1967 GGAACCAACAGCTGTCTCAGAGGCCAGATCTCCATTTTATGATGTAGAAATGCCCCAAGG 2026
QY 2249 ATAACCCCATGTACTACTGTATAACAGAAAGACAGCTGCAGAGCTGTGCCCTGGAACAGA 2308
DB 2027 ATAACCCCATGTACTACTGTATAACAGAAAGACAGCTGCAGAGCTGTGCCCTGGAACAGA 2086
QY 2309 ACTGCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGCA 2368
DB 2087 ACTGCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGCA 2146
QY 2369 TTGGCTGCAATTTGGTTGGAAAACCTCATGTTTGAATAATTAATCTGCTGCAAGAGAAATATG 2428
DB 2147 TTGGCTGCAATTTGGTTGGAAAACCTCATGTTTGAATAATTAATCTGCTGCAAGAGAAATATG 2206
QY 2429 ACAATGCTAAATTTGTTGTAGGAAACCAATGCCCCTTTGGCTTCTCTTCAACCCAGA 2488
DB 2207 ACAATGCTAAATTTGTTGTAGGAAACCAATGCCCCTTTGGCTTCTCTTCAACCCAGA 2266
QY 2489 AGAAGTGAATTTGCTCTTAAGCAGCTGCGAATAATGAGTCAATCTCAGAGCATGTCCA 2548
DB 2267 AGAAGTGAATTTGCTCTTAAGCAGCTGCGAATAATGAGTCAATCTCAGAGCATGTCCA 2326
QY 2549 AGCTCACTTAAACCCCATGGTTCGCCCTTCGGAAGATCAATGTGCTTCTTACTGCTGGG 2608
DB 2327 AGCTCACTTAAACCCCATGGTTCGCCCTTCGGAAGATCAATGTGCTTCTTACTGCTGGG 2386
QY 2609 AAGATATGTCCTTCAATTAATAGTCTTACTACAGTGGATGCGCTCAGGCCAGTGTG 2668
DB 2387 AAGATATGTCCTTCAATTAATAGTCTTACTACAGTGGATGCGCTCAGGCCAGTGTG 2446
QY 2669 CTGGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728
DB 2447 CTGGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506
QY 2729 TCACCCCATCTAATGGTGTGCTGTGAAGGCCCTGCAAAACACAGTGTGTAAGCAGTGGC 2788
DB 2507 TCACCCCATCTAATGGTGTGCTGTGAAGGCCCTGCAAAACACAGTGTGTAAGCAGTGGC 2566
QY 2789 GGACACCATGTCCCTTTGAGGACAGCATGTGGAGATTTGCACACAGCGGCACTGTGAGTGA 2848
DB 2567 GGACACCATGTCCCTTTGAGGACAGCATGTGGAGATTTGCACACAGCGGCACTGTGAGTGA 2626
QY 2849 TGTGGTGCAGCAACATGAAGCAGTGTGGAATCCAAATGCTTATGTGGCTCTCTTCCCTT 2908
DB 2627 TGTGGTGCAGCAACATGAAGCAGTGTGGAATCCAAATGCTTATGTGGCTCTCTTCCCTT 2686
QY 2909 TTGGCCAGTGTATGGATGGTATAGATGAGCACCTGCCCTGAAATTTGTTTCTCAGGCT 2968
DB 2687 TTGGCCAGTGTATGGATGGTATAGATGAGCACCTGCCCTGAAATTTGTTTCTCAGGCT 2746
QY 2969 ACTGTACCTGTAGTCAATTTCTTTGGAGCAACCAAGGCTGTGGCTGTGTAATGATCCAGCA 3028
DB 2747 ACTGTACCTGTAGTCAATTTCTTTGGAGCAACCAAGGCTGTGGCTGTGTAATGATCCAGCA 2806
QY 3029 ATACTGGCAAGGAAAATCCATAGAGGTTCCCTATAAAGGACAGTGAAGTGCCTTCCG 3088
DB 2807 ATACTGGCAAGGAAAATCCATAGAGGTTCCCTATAAAGGACAGTGAAGTGCCTTCCG 2866
QY 3089 AAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGG 3148
DB 2867 AAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGG 2926
QY 3149 ACAGCAGATPACAACTGGTCTTTTCATTTCACTGTCCAGCTTGGCCAAATGCAACGGCCACAGTA 3208
DB 2927 ACAGCAGATPACAACTGGTCTTTTCATTTCACTGTCCAGCTTGGCCAAATGCAACGGCCACAGTA 2986
QY 3209 AATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTCTGACCACTGACCAAGCAAGCTGCG 3268

DB 2987 AATGATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACACAGGAGAGCATGGG 3046
 QY 3269 AGACCTGATATCTGGCTTCTAOGGTGATCCCAACATGGAGGAAATGTGAGCCATGCA 3328
 DB 3047 AGACCTGATATCTGGCTTCTAOGGTGATCCCAACATGGAGGAAATGTGAGCCATGCA 3106
 QY 3329 AGTGAATGGGACGCTCTGTGACACACACAGGCAAGTGTCTGACACACCA 3388
 DB 3107 AGTGAATGGGACGCTCTGTGACACACACAGGCAAGTGTCTGACACACCA 3166
 QY 3389 AGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTGAAGAAATCGATACCAAGGAACC 3448
 DB 3167 AGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTGAAGAAATCGATACCAAGGAACC 3226
 QY 3449 CTCTCAGAGGACATGTTATTATATCTCTTATGACTATCAGTTCACCTTTAGTCTAT 3508
 DB 3227 CTCTCAGAGGACATGTTATTATATCTCTTATGACTATCAGTTCACCTTTAGTCTAT 3286
 QY 3509 CCCAGGAAGATGATCGCTATTATACACAGCTATCAATTTTGGCTACTCCTGACGACAAA 3568
 DB 3287 CCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGGCTACTCCTGACGACAAA 3346
 QY 3569 ACAGGATTTGGACATGTTATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGG 3628
 DB 3347 ACAGGATTTGGACATGTTATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGG 3406
 QY 3629 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3688
 DB 3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3466
 QY 3689 ACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAAAGTTTGATTTTCGCAACCCACCAA 3748
 DB 3467 ACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAAAGTTTGATTTTCGCAACCCACCAA 3526
 QY 3749 ATATCAGTTTCTGTTTATGTCAGTAATTTACCTGCGCCCATCAAAATTCAGAT 3803
 DB 3527 ATATCAGTTTCTGTTTATGTCAGTAATTTACCTGCGCCCATCAAAATTCAGAT 3581

RESULT 10

ADD70997
 ID ADD70997 standard; DNA; 3597 BP.
 XX
 AC ADD70997;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human attractin gene SEQ ID NO:1.
 XX
 KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
 KW cytosolic; gene therapy; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003061564-A2.
 XX
 PD 31-JUL-2003.
 XX
 PF 20-DEC-2002; 2002WO-US040718.
 XX
 PR 21-DEC-2001; 2001US-0341815P.
 PR 31-DEC-2001; 2001US-0343185P.
 XX
 PA (GENE-) GENE LOGIC INC.
 PA (LGBI-) LG BIOMEDICAL INST.
 XX
 PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;
 XX
 DR WPI; 2003-663343/62.
 XX
 PT Diagnosing liver cancer cells, useful for treating liver cancer
 PT associated with chronic hepatitis or cirrhosis comprises detecting the

PT Level of expression in a tissue sample of one or more genes associated
 PT with cancerous liver tissues.

Claim 1; SEQ ID NO 1; 176pp; English.

XX The present invention describes a method for diagnosing liver cancer
 XX cells comprising detecting the level of expression in a tissue sample of
 CC one or more genes given in the specification (see ADD70997 to ADD71105),
 CC where differential expression of the genes is indicative of liver cancer.
 CC Also described: (1) detecting the progression of liver cancer in a
 CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)
 CC patient; (5) detecting the presence or progression of liver cancer in a
 CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver
 CC cancer related to chronic hepatitis from liver cancer related to
 CC cirrhosis; (7) screening for an agent capable of modulating the onset or
 CC progression of liver cancer; (8) a composition comprising at least two
 CC oligonucleotides comprising a sequence that specifically hybridizes to
 CC any of the genes; (9) a solid support comprising the at least two
 CC oligonucleotides; (10) a computer system comprising a database containing
 CC information identifying the level in liver tissue of a set of genes; (11)
 CC a method for using the computer system to present information identifying
 CC the expression level in tissue or cell of any of the genes; and (12) a
 CC therapeutic agent for slowing or halting the progression of liver cancer.
 CC The methods are useful for treating liver cancer associated with chronic
 CC hepatitis or cirrhosis. The present sequence represents a specifically
 CC claimed human gene sequence which is used in the exemplification of the
 CC present invention.

XX Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 80.9%; Score 3469.4; DB 9; Length 3597;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 3507; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

QY 270 GTCCGGCTCAGCCGCGAGGCGCAAGATGTGACCGGCCCTGTGTCAACGGGGTCG 329
 DB 48 GCGCGCGAGCGGAGCGCTCGCGGGGAGGCGGGCGGCGGCGCTGTGTCAACGGGGTCG 107
 QY 330 CTGCAACCCCTGGCACCGGCGAGTGGCTCTGCCCGCGCGCTGGGTGGCGGAGCAATGCCA 389
 DB 108 CTGCAACCCCTGGCACCGGCGAGTGGCTCTGCCCGCGCGCTGGGTGGCGGAGCAATGCCA 167
 QY 390 GCATCGCGGGGCGCTTCAGACTAACTGGATCTCTCGGGTTTGTGACAGATGACCTGG 449
 DB 168 GCATCGCGGGGCGCTTCAGACTAACTGGATCTCTCGGGTTTGTGACAGATGACCTGG 227
 QY 450 AAATTATATAATACAAACGAAAGTGCACGTGCTCATTTGAAGGACAGCCAAATAGATAAT 509
 DB 228 AAATTATATAATACAAACGAAAGTGCACGTGCTCATTTGAAGGACAGCCAAATAGATAAT 287
 QY 510 GAGACTTCGTTTCAATCAATTTTGTGACAGAGTGTAGTTGGGACCATTTATATGTTTGA 569
 DB 288 GAGACTTCGTTTCAATCAATTTTGTGACAGAGTGTAGTTGGGACCATTTATATGTTTGA 347
 QY 570 TGGGAGCTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCTCATTTGTTCTGAGAG 629
 DB 348 TGGGAGCTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCTCATTTGTTCTGAGAG 407
 QY 630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGGCCACATCAGGTTATGCTTGTGCAATTT 689
 DB 408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGGCCACATCAGGTTATGCTTGTGCAATTT 467
 QY 690 TTTTAGTGTGCTGTATATAATTTGACTGGATTTAAATATTACTTACAGTTTGTATGTG 749
 DB 468 TTTTAGTGTGCTGTATATAATTTGACTGGATTTAAATATTACTTACAGTTTGTATGTG 527
 QY 750 TCCAAATTAACCTGCTCAGGCGGAGGAGAGTGTAAAGTACAGTAATAGCAGGCAAACTGTTGA 809
 DB 528 TCCAAATTAACCTGCTCAGGCGGAGGAGAGTGTAAAGTACAGTAATAGCAGGCAAACTGTTGA 587
 QY 810 ATGTGAATGTTCTGAAATCTGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Db	588	ATGTGAATGTTCTGAAAACTGGAAAAAGGTGAAGCATGTGACATTCCTCTACCTGTACAGACAA	547
Qy	870	CTGTGGTTTTTCCTCATCGAGGCACTCGCAATTCAAGTGATGTGAGAGGATGCTCTCGTT	929
Db	648	CTGTGGTTTTTCCTCATCGAGGCACTCGCAATTCAAGTGATGTGAGAGGATGCTCTCGTT	707
Qy	930	CTCAGACTGGCAGGGTCCCTGGATGTTTCAGTTCCTGTACCAAGCTAAACAGTCAATTTGGAC	989
Db	708	CTCAGACTGGCAGGGTCCCTGGATGTTTCAGTTCCTGTACCAAGCTAAACAGTCAATTTGGAC	767
Qy	990	TCGAGAGGAATATTCCTAAACTTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGTCAATGG	1049
Db	768	TCGAGAGGAATATTCCTAAACTTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGTCAATGG	827
Qy	1050	AAACATTAATGTGGGTTGTTGGAGGATATATGTTTCACCACTCAGATATTAACATGTTCT	1109
Db	828	AAACATTAATGTGGGTTGTTGGAGGATATATGTTTCACCACTCAGATATTAACATGTTCT	887
Qy	1110	AGCGTAGACTTGTCTCTAGGAGTGGCTTCACATAAACCGTTCGTGTAAACAATGTGGT	1169
Db	888	AGCGTAGACTTGTCTCTAGGAGTGGCTTCACATAAACCGTTCGTGTAAACAATGTGGT	947
Qy	1170	TGTTAGATATGGTCAATCTTTTGGCATATATACAAGGATATAAATTTACATGTGAGGAAA	1229
Db	948	TGTTAGATATGGTCAATCTTTTGGCATATATACAAGGATATAAATTTACATGTGAGGAAA	1007
Qy	1230	AATTTGATCACTGGGAATGTGACCAATGAGTTTGAGAGTTTTTTCATTCATTAATGAGTC	1289
Db	1008	AATTTGATCACTGGGAATGTGACCAATGAGTTTGAGAGTTTTTTCATTCATTAATGAGTC	1067
Qy	1290	ATGGGTGTTCTTGACCCCTAAGGCAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA	1349
Db	1068	ATGGGTGTTCTTGACCCCTAAGGCAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA	1127
Qy	1350	CATTGTTACTGAGAAATGGCCGAGTGGTTCATGCTGGTTCATCTTTGGTCACTGCCCTCT	1409
Db	1128	CATTGTTACTGAGAAATGGCCGAGTGGTTCATGCTGGTTCATCTTTGGTCACTGCCCTCT	1187
Qy	1410	CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT	1469
Db	1188	CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT	1247
Qy	1470	ACACACCCAGGTGCCCTTGTCGAAGGGGTTACGGCCATGACAGTGTTTACGACCATAG	1529
Db	1248	ACACACCCAGGTGCCCTTGTCGAAGGGGTTACGGCCATGACAGTGTTTACGACCATAG	1307
Qy	1530	GACAGGGCCCTATACGTTTCATGTGTGCTACAAGGCTTTTCAGTGCCAAATAGTACCGCT	1589
Db	1308	GACCAGGGCCCTATACGTTTCATGTGTGCTACAAGGCTTTTCAGTGCCAAATAGTACCGCT	1367
Qy	1590	TGCAGATGATCTCTACCGATATGATGTGGATACCAGATGTGGACCAATCTTAAGGACAG	1649
Db	1368	TGCAGATGATCTCTACCGATATGATGTGGATACCAGATGTGGACCAATCTTAAGGACAG	1427
Qy	1650	CCGATTTTTCCGTTTACTTGGACACACGCTGTGATAGTGGGAACCATGCTGTGTTGG	1709
Db	1428	CCGATTTTTCCGTTTACTTGGACACACGCTGTGATAGTGGGAACCATGCTGTGTTGG	1487
Qy	1710	GGGAAACACACAAATGACACATCTATGAGCCATGGGCGCAAAATGCTTCTTTCAGATTT	1769
Db	1488	GGGAAACACACAAATGACACATCTATGAGCCATGGGCGCAAAATGCTTCTTTCAGATTT	1547
Qy	1770	CATGGCCTATGACATTCGCTGTGACCGGTGTGATGCTTCCGAGACCTGATCTCCACCA	1829
Db	1548	CATGGCCTATGACATTCGCTGTGACCGGTGTGATGCTTCCGAGACCTGATCTCCACCA	1606
Qy	1830	TGATGTCAACAGATTTGGCCATTTCAGAGTCTTTACACAACAGCACCATGTATGTGTTCG	1888
Db	1607	TGATGTCAACAGATTTGGCCATTTCAGAGTCTTTACACAACAGCACCATGTATGTGTTCG	1666
Qy	1889	GTGGTTTTCAATPAGTCTCCTCTCAGCGACATCTCTGGTATTCACTCGGAAACAGTGTGATG	1948
Db	1667	GTGGTTTTCAATPAGTCTCCTCTCAGCGACATCTCTGGTATTCACTCGGAAACAGTGTGATG	1726

Db 493 ACGCTTAATACAGCTGTTCGGAAGGTGAGGCTCGAGGAACAGTTCGAGGCAAGCTTC 552
Qy 411 ----- 410
Db 553 GGCTACAGAAATAAGTTCAAGAGTAACCTCGGGCAAATTGGGCTGTCTCCAAACCAAAA 612
Qy 411 ----- 410
Db 613 TGAGCGAAAAGGACAGCTAGAGTCTTTTGGGAAAATTTAGCTGACTAAATTTTTCACC 672
Qy 411 ----- ACTAATCTGATCTTCTGGGTTTGTGACAGATGGAACCTGGAAATTAATAATACAAA 466
Db 673 GAGAACTAACTGGCTCTTCTGGATTTGTAAACAGATGGACCTGGGAATTAATAATAAAGA 732
Qy 467 CGAAGTSCAGTGCCTCATTAAGAGGACAGCAATAAGATAATAGAGATTCGTTTCAATC 526
Db 733 CGAAGTSCAGATGCTCATTTAAGAGGACGCAATAGATTAATAGAGCTTCGTTCAACC 792
Qy 527 ATTTTGTACAGAGTGTAGTTGGACCAATTAATATATGTTTATGATGGGACTCAATTTATG 586
Db 793 ATTTTGTACAGAAATGTAGCTGGACCAATTAATATATGTTTATGATGGGACTCAATCTAG 852
Qy 587 CACCGTAGTGTGCTGCAATTTAGTGGCTCAATGTTCTGAGAGAGATGCAATGAGACTG 646
Db 853 CACCTGTGATGCTGCTTTAGTGGCTCAATGTTCTGAAAGAGATGGCAATGAGACGG 912
Qy 647 TCCCTGAGGTTGTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTT 706
Db 913 CTCTGAGGTCACGTGCTCAGTTATGCACTGCTGCAATTTTTCAGTATGCTGCTT 972
Qy 707 ATAAATTTGACTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 766
Db 973 ATAAATCTGACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1032
Qy 767 GCCGAGAGAGTGTAAGATCAGTAATAGCAGGAACTGTTGAATGTAATGTTCTGAAA 826
Db 1033 GCCGAGAGAGTGTAAGAGAGTAAGCAGAGCGCTGTTGAGTGTGAATGTTCTGAAA 1092
Qy 827 ACTGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAATGTGGTTCCTCATC 886
Db 1093 ACTGAAAGGGAGTCTGTGACATTCCTCACTGTACAGACAATGTGGTTCCTCAAC 1152
Qy 887 GAGCANTCTGAATTCAGTGTGACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 946
Db 1153 GAGCANTCTGTAATGCAAGGATACCAAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1212
Qy 947 CTGATGTTCACTTCCTGTACAGCTAACCACTCAATTTTGGACTCGAGAGGAATATTCTA 1006
Db 1213 CTGATGTTCAATTCCTGTGACAGTAACCACTCTTTTGGACTCGAGAGGAATATTCTG 1272
Qy 1007 ACTTAAAGCTCCGAGAGCATCTCAATAGCTGTGTCATGAGGAACATTAATGAGGTTG 1066
Db 1273 ATTTAAAGCTTCCGAGAGCTCTCATAAAGCTGTGGTCAATGGAATTAATGAGGTTG 1332
Qy 1067 TTGAGGATATATGTTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTGCTT 1126
Db 1333 TTGCGGATATATGTTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTGACTT 1392
Qy 1127 CTAGGAGTGGCTTCCATTAACCGTTCTGTGAAACAATGTTGTTTGTATGATGGTCAAT 1186
Db 1393 CTAGGAGTGGCTTCCATTAACCACTCTGTGAAACAATGTTGTTTGTATGATGGTCAAT 1452
Qy 1187 CTGTTGGCATTAACAGGATTAATTTACATGATGAGGAAATTAATGATCCACTGGGA 1246
Db 1453 CTGTTGGCATTAACAGGATTAATTTACATGATGAGGAAATTAATGATCCACTGGGA 1512
Qy 1247 ATGTGACCAATGAGTTGAGAGTTTTCATATTCATATGATGATGATGATGATGATGAT 1306
Db 1513 ACGTGACCAATGAGCTGAGAGTATTTTCAATTCATATGATGATGATGATGATGATGAT 1572
Qy 1307 CTAAGGCAAGAGAGATGAGTGTGGGCACTCTGCAACATTTGTTACATCTGAAGA 1366

Db 1573 CGAAAGCTTAAGGATCAGTATGCAAGTGTGGACACTCAGCACACATTTGTTACCTGGCAT 1632
Qy 1367 ATGGCCGAGTGTGATGCTGCTCATCTTTGTTGCTCACTGCGCTCTCTATGATATATAAGCA 1426
Db 1633 CTGGCCGTGTGATGCTGCTCATCTTCGCTCATCTTCGCTCATCTATGATATATAAGCG 1692
Qy 1427 ATGTGACAGAAATATGATTTGATTAAGAAACATGAGGATATATTACACACCCAGGTCGCC 1486
Db 1693 TTGTGACAGAAATATGATTTGAAAGAAACACATGAGGATATATTACATACTCAGGTTGCTC 1752
Qy 1487 TTGTGCAAGGGGTTTACGGCCATAGCAGTGTTTACGACCATAGGACACAGGGCCCTATACG 1546
Db 1753 TTGTGCAAGGGGTTTATGCCCCACAGTAGTGTTTATGATGACAGACCAAGGCTCTGACG 1812
Qy 1547 TTCTATGTTGGCTTACAAAGCTTTTCAAGTGTGCAATAAGTACCGGCTTGAGATGATCTTACC 1606
Db 1813 TTCTATGTTGGCTTACAAAGCTTTTCAAGGCTTTCAAGGCAACAAATACCGGCTTGACATGACCTCTACA 1872
Qy 1607 GATATGATGTGGATACCCAGATGTGGACCATTTCTTAAGCAGACCGGATTTTTCGTTACT 1666
Db 1873 GATAGATGTGGATGACTCAGATGTGGACCATTTCTTAAGCAGACCGGATTTTTCGTTACT 1932
Qy 1667 TGACACACAGCTGTGATGATGAGTGGAAACCATGCTGTGTTTGGGGGAAACACACAATG 1726
Db 1933 TGCAATACAGCTGTGATGATGAGTGGAAACCATGCTGTGTTTGGAGGAAACACACAATG 1992
Qy 1727 ACACATCTATGAGCCATGCGGCCAATGCTTCTCTTCAAGATTTTCAATGCTTATGACATG 1786
Db 1993 ACACATCTTCCATGAGCCATGCGGCCAATGCTTCTCTCTCGGACTTCAATGCTTATGACATG 2052
Qy 1787 CCGTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATCTCAACAGATTTG 1846
Db 2053 CTTGTGACCGATGCTGCTGCTTCCAGACCTGAGCTCCAATCATGATGATCAACAGATTTG 2112
Qy 1847 GCCATTCAGAGCTTTTACACACAGACCATGATGTGTCGTTGTTTCAATGATCTTCC 1906
Db 2113 GCCATTCAGAGCTTTTACACACAGACCATGATGTGTCGTTGTTTCAACAGCTTCC 2172
Qy 1907 TCCTCAGGACATCTGTTATTCCTCGGAACAGTGTGATGTCGCTCGGATCGAGTGAAGCG 1966
Db 2173 TCCTCAGTACCTTTGTTTACCTCGGACATGATGTGTCGTTGTTTCAACAGCTTCC 2232
Qy 1967 CTTGTTTACAGCAGGACCTGTTATTCGTTGTTGTTGAAACACAGGTCGTCTCAGTGT 2026
Db 2233 CTTGTTGTCAGCAGGACCTGTTATTCGTTGTTGTTGAAACACAGGTCGTCTCAGTGT 2292
Qy 2027 TCTGTTGGCGCTGCGCACTGATGAACAGAAAGTTAAATCAGAACTGTTTTTCCA 2086
Db 2293 CTTCTGCGAGTGGCACTGAAAGCAAGCAGAAAGTTAAATCAGAGTGTTTTTCTA 2352
Qy 2087 AAAGAACTCTTCAACATGACAGATGTACACAGACACAGATTTGTTACAGCTGTACAGCCA 2146
Db 2353 AAAGAACTCTTCAACATGACAGATGTACACAGACACAGATTTGTTACAGCTGTACAGCCA 2412
Qy 2147 ACACCAATCACTGCGCACTGGTGAATGACCATGTTGTTCCCGAGGAACCAAGCTGCTAG 2206
Db 2413 ATACCAATCACTGCGCACTGGTGAATGACCATGTTGTTCCCGAGGAACCAAGCTGCTAG 2472
Qy 2207 AAGGCCAGATCTCCATTTTTTAGGTATGAAATTTGCCCAAGGATTAACCCCATGTACTACT 2266
Db 2473 AAGGCCAGATCTCCATTTGCAAGTATGAGTTGCCCCAAGGATTAACCCCATGTACTACT 2532
Qy 2267 GTAAACAGAGAGACCACTGCGAGGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCC 2326
Db 2533 GCAATTAAGAAACCACTGCGAGGCTGTGCCCTAGACACCAAGCTGCCAGTGGAGCCCC 2592
Qy 2327 GGAATCAGAGTGCATTTGCCCGGAAAATATCTGTGGCATTTGGCTGCAATTTGGTTG 2386
Db 2593 GGAATCAGAGTGCATTTGCCCGGAAAATATCTGTGGCAATTTGGCTGCAATTTGGTTG 2652
Qy 2387 GAAATCTATGTTTGAATAATTAATCTGCTGCAAGAGAAATTAATGACATGCTAAATTTTCT 2446
Db 2653 GAAATCTGCTGCTGAATAATTAATCTGCTGCAAGAGAAATTAATGACATGCTAAATTTTCT 2712

QY 2447 GTAGGAACCAATGCGCTTTTGGCTTCTCTTACAAACCAGAGAGGTAGAAATTTGTC 2506
Db 2713 GTAGGAACCAATGCGCTTTTGGCTTCTCTTACAAACCAGAGAGGTAGAAATTTGTC 2772
QY 2507 TTAAGCAGCTGCGAATAATGAGTCACTCAGAGCATGTCCAAAGCTCACCTTAACCCCAT 2566
Db 2773 TTAAGCAGCTGCGAATAATGAGTCACTCAGAGCATGTCCAAAGCTCACCTGACTCCAT 2832
QY 2567 GGGTCGGCTTCGGAAGATCAATGTCTCTACTGTCTGCTGGAAGATATGTCCCATTTA 2626
Db 2833 GGGTGGCTTCGGAAGATCAATGTCTCTACTGTCTGCTGGAAGATATGTCTCCATTC 2892
QY 2627 CAATAGTTTACTACAGTGGATGCGCTGAGCCAGTGTCTGGAATTTCTGTGGAATTT 2686
Db 2893 CAATAGTTTCTGAGTGGATGCCATCTGAGCCAGTGTCTGGAATTTCTGTGGAATTT 2952
QY 2687 TATCAGAAACCCAGTACTCGGGAGTGAAGGCTGCAACTGTCATCAACCCACTCAATGGTA 2746
Db 2953 TGTGAGAGCTTAGTACTCGGGATTTAAAGGCTGCAACTGTCATCAACCCCTCTCAATGGCA 3012
QY 2747 GTGTCTGGAAGGCTGCAACACACAGTGTCTGAGTGTGCGGACACCATGTGCTTGA 2806
Db 3013 GCGTCTGGAAGGCTGCAACACACAGTGTCTGAGTGTGCGGACACCATGTGCTTGA 3072
QY 2807 GGACAGCATGTGGAGATTGCAACAGCGGACGCTCTGAGTGTGCGGACACCATGTG 2866
Db 3073 GGACAGCATGTGGAGATTGCAACAGCGGACGCTCTGAGTGTGCGGACACCATGTG 3132
QY 2867 AGCAGTGTGAGTCTCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTGGAAT 2926
Db 3133 AGCAGTGTGAGTCTCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTGGAAT 3192
QY 2927 GGTATACCATGAGCACTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAAT 2986
Db 3193 GGTATACCATGAGCACTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAAT 3252
QY 2987 GCTTGGAGCAACAGAGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3046
Db 3253 GCTTGGAGCAGCAGGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3312
QY 3047 GCATAGAGGTTCTTATAAGAGCAGTGAAGATGCTTCCGAGCCCTTACAGGAAT 3106
Db 3313 GTATTGAGGAGCAGTATAAGAGCAGTGAAGATGCTTCCGAGCCCTTACAGGAAT 3372
QY 3107 TCTATCCAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATCAACTGGT 3166
Db 3373 TGTATCCAGCCCTTCTGAATCTCCAGCATGTGTCTAGAGGACAGCAGATCAACTGGT 3432
QY 3167 CTTTCACTGCTCCAGCTGCAATGCAACGCGCACAGTAAATGCATCAATCAGAGCA 3226
Db 3433 CTTTCACTGCTCCAGCTGCAATGCAACGCGCACAGTAAATGCATCAATCAGAGTA 3492
QY 3227 TCTGTGAGAGTGTGAGAACTGACCAAGCAGCACTGCGAGACTGTCATATCTGGCT 3286
Db 3493 TCTGTGAGAGTGTGAGAACTGACCAAGCAGCACTGCGAGACTGTCATATCTGGCT 3552
QY 3287 TCTACGCTGATCCCAATGAGGGAATGTGAGCCATGCAATGAGGACGCTG 3346
Db 3553 TCTATGAGTACCCGACTTAATGAGGGAATGTGAGCCATGCAATGAGGACGCTG 3612
QY 3347 CTCCTGCAACACCAACAGCGGAGTGTCTTCTGACCAACAGGCGCTCAAGGGGACG 3406
Db 3613 CACTGTGCAACACCAACAGCGGAGTGTCTTCTGTACCAACAGGCTCAAGGGGACG 3672
QY 3407 AGTCCAGCTATGTAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTT 3466
Db 3673 AGTCCAGCTATGTAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTT 3732
QY 3467 ATTATACCTCTTATTGATATCAGTTTCACTTTTATCTATCCAGGAGATGATCGCT 3526
Db 3733 ACTATACCTCTTCACTTATCAGTTTCACTTTTATCTATCCAGGAGACGACCGCT 3792

QY 3527 ATTACACAGCTATCAATTTTGGCTACTCTCTGACGAAACAAACAGGATTTGGCATGT 3586
Db 3793 ACTACACAGCTATCAATTTTGGCTACTCTCTGATGAACAAACAGGATTTGGCATGT 3852
QY 3587 TCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTG 3646
Db 3853 TCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCCAGCG 3912
QY 3647 GAACCCAGCTCGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGAGATCAAAAG 3706
Db 3913 GAACCCAGCTCGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGAGATCAAAAG 3972
QY 3707 ATAGTTTCTCTAATGAGAAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTT 3766
Db 3973 ATAGTTTCTCTAATGAGAAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTT 4032
QY 3767 ATGTCAATGATTTTCACTGGCCATCAAAATTCAGATTCCTTCTCAGCACAGCAAT 3826
Db 4033 ATGTCAATGATTTTCACTGGCCATCAAAATTCAGATTCCTTCTCAGCACAGCAAT 4092
QY 3827 TTATGACCTGTGACAGTCTTCTGCTGACTTTCTTCAGTTGTTTCTCTCTTCTCTGG 3886
Db 4093 TCATGACCTGTGACAGTCTTCTGCTGACTTTCTTCAGTTGTTTCTCTCTCTCTCTGG 4152
QY 3887 TGGCTGCTGTGTTTGGAGATCAAAACAAAGTTGTTGGGCTTCCAGAGCTAGAGAGCAAC 3946
Db 4153 TGGCTGCTGTGTTTGGAGATCAAAACAAAGTTGTTGGGCTTCCAGAGCTAGAGAGCAAC 4212
QY 3947 TTCTTTCGAGAGTGAACAGATGGCCAGCGCTTCCCTCTGCTCTGTAATGTCGCTTGG 4006
Db 4213 TTCTTTCGAGAGTGAACAGATGGCCAGCGCTTCCCTCTGCTCTGTAATGTCGCTTGG 4272
QY 4007 AAACAGATGAGAGGCTCTCTGATCTTATTGGGGGAGTATAAAGACTGTTCCCAACCCCA 4066
Db 4273 AAACAGATGAGAGGCTCTCTGATCTTATTGGGGGAGTATAAAGACTGTTCCCAACCCCA 4332
QY 4067 TTGCACTGGAGCCGCTGTTTGGCAACAAAGCGCTGCTCTCTGTTGTTGAGGCTCC 4126
Db 4333 TTGCACTGGAGCCGCTGTTTGGCAACAAAGCGCTGCTCTCTGTTGTTGAGGCTCC 4392
QY 4127 CTCGAGGCTGCTGGCATCCCTCTCTGCGCAGTCAAGTCTTCTGCTGCGCAGCGCC 4186
Db 4393 CTCGAGGCTGCTGGCATCCCTCTCTGCGCAGTCAAGTCTTCTGCTGCGCAGCGCC 4452
QY 4187 TGGTGGACATTTCTCAGCAGATGCGATAGTGTAAAGGAGAGTCAAGGCTGAGAA 4246
Db 4453 TGGTGGACATTTCTCAGCAGATGCGATAGTGTAAAGGAGAGTCAAGGCTGAGAA 4512
QY 4247 ACCGGAAGCAGCGCCCTTCCAGCAGCTGGAGCTGCTCTGA 4290
Db 4513 ACCGGAAGCAGCGCCCTTCCAGCAGCTGGAGCTGCTCTGA 4556

RESULT 12

AAS72660
ID AAS72660 standard; cDNA; 3490 BP.

XX AAS72660;

AC AAS72660;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #8464.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

[illegible]

1906 CTCCTCAGCGACATCCTGGTATTACCTCGGAACAGTGTGATGGCATCGGAGTGAAGCC 1965
1522 CTCCTCAGCGACATCCTGGTATTACCTCGGAACAGTGTGATGGCATCGGAGTGAAGCC 1581
1966 GCTTGTATTAGCAGAGGACCTGGTATTCTGGTGTGTGTGGAAACAGAGGTCTCTCAGTGT 2025
1582 GCTTGTATTAGCAGAGGACCTGGTATTCTGGTGTGTGTGGAAACAGAGGTCTCTCAGTGT 1614
2026 ATCTCGTGGGCGCTGGCAACTGATGAACAAGAGAAAGTTAAATCAGAGATGTTTTTCC 2085
1615 ----- 1614
2086 AAAAGAACTCTTGACCATGACAGATGTGACAGACACAGATTTTACAGCTGTACAGCC 2145
1615 -----CTCTTTGACCATGACAGATGTGACAGACACAGATTTTACAGCTGTACAGCC 1667
2146 AACCCATGACTGCCACTGGTGAATGCAATGATGTGTCCTCCAGGACACCAAGCTGTCTCA 2205
1668 AACCCATGACTGCCACTGGTGAATGCAATGATGTGTCCTCCAGGACACCAAGCTGTCTCA 1727
2206 GAAGGCCAGATCTCCATTTTATAGTATGAGATTTGCCCAAGGATAACCCCATCTACTAC 2265
1728 GAAGGCC-----AGGTATGAGAAATTGCCCAAGGATAACCCCATCTACTAC 1773
2266 TGTAAACAGAGACCCAGCTGACAGAGCTGTGCTCTGACACAGAACTGCCAGTGGAGCCC 2325
1774 TGTAAACAGAGACCCAGCTGACAGAGCTGTGCTCTGACACAGAACTGCCAGTGGAGCCC 1833
2326 CGGAATCAGGAGTGCAATGCTCCCGAAATATCTGTGCAATGCTGCGCATTTTGGTT 2385
1834 CGGAATCAGGAGTGCAATGCTCCCGAAATATCTGTGCAATGCTGCGCATTTTGGTT 1893
2386 GGAACTCATGTTGAAATTTACTATGCTGCAAGGAGAAATATGAAATGCTAAATTTGTT 2445
1894 GGAACTCATGTTGAAATTTACTATGCTGCAAGGAGAAATATGAAATGCTAAATTTGTT 1953
2446 TGTAGGACCAATGCCCTTTTGGCTCTCTTACAAACCCAGAGAGGTA-GAATTTGT 2504
1954 TGTAGGACCAATGCCCTTTTGGCTCTCTTACAAACCCAGAGAGGTA-GAATTTGT 2013
2505 CTTTAAAG--CAGCTGCGAATATGCA-GTATCTCAGA-GCATGTCCAA-GCTCACCTTA 2559
2014 CTTTAAAGCAGCTGCGAATATGCAAGTGTATCTCAGAGGATGTCCAAAGCTCACCTTA 2073
2560 ACCCATGGG---TCGGCTTCGGAAGATCAA-TGTGTCTACTGTGTGGGAGATAT 2615
2074 ACCCATGGGTCGGGCTTCGGGAAAGATCAATTTGTCTCTACTGTGTGGGAGATAT 2133
2616 GTCCCCCATTTACAAATAGTTTACTACAGTGTGATGCCCTCTGAGCCAGTGTGATGCTGAT 2675
2134 GTCCCCCATTTACAAATAGTTTACTACAGTGTGATGCCCTCTGAGCCAGTGTGATGCTGAT 2193
2676 CTGTGGAAATTT-----ATCAGAACCCAGTACTCGGGAGCTGAAGGTGCAACCTGCAAT- 2729
2194 CTGTGGAAATTTTATTTCAGGAACCCAGTACTTCGGGGAGCTGAAGGTGCAACCTGCAAT 2253
2730 -CAACCCACTCAATGGTGTGTGTGTGAAA--GGCTGTGAAACCCAGT-GCTAAGCAGT 2785
2254 CAACCCACTCAATGGTGTGTGTGTGAAAAGGCTGTGCAACCCAGTGTGATGCTGAT 2313
2786 GCCGGACCAATGTGCC-----TTGAGGACAGCATGTGAGATTTGCCACCGCAGCTCT 2841
2314 GCCGGACCAATGTGTGCCCTTTGAGGACAGCATGTGAGATTTGCCACCGCAGCTCT 2373
2842 GAGT-GCATGTGTGCGACACACAGAA-GCAGTGTGTGGACTCCAATGCTATGTGGCT 2899
2374 GAGTGGCATGTGTGCGACACACAGAAAGGAGTGTGTGGACTCCAATGCTATGTGGCT 2433
2900 CTTTCCCTTTT-GGCCAGTGTATGGAATGTTATACGATGA-GCAGCTGCCCCCTGAAAA 2957
2434 CTTTCCCTTTTGGGCCAGTGTATGGAATGTTATACGATGAGGACCTGTGCCCTTAAAA 2493
2958 TTGTTCCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACCA-GGCTGTGGCTGTGTGA 3016

2494 TTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACACAGGCTGTGGCTGTGA 2553
3017 CTGATCCAGCAATACTGCAAGGGAATGATAGAGGGTTCCTATTAAGGACCACTGA 3076
2554 CTGATCCAGCAATACTGCAAGGGAATGATAGAGGGTTCCTATTAAGGACCACTGA 2613
3077 AGATGCTTTGCAAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATTCAGCA 3136
2614 AGATGCTTTGCAAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATTCAGCA 2673
3137 TGTGCTAGAGGACAGAGATCAACTGCG--TCTTTCTACTGTCAGGCTTGCATATG 3194
2674 TGTGCTAGAGGACAGAGATCAACTGCGGCTCTTTCTACTGTCAGGCTTGCATATG 2733
3195 CAAC-GGCCACACAGTAAATGCAATCAGAGATCTGTGAGAAAGTGTGAGAACCTGACCA 3253
2734 GAAAGGACACAGTAAATGCAATCAGAGATCTGTGAGAAAGTGTGAGAACCTGACCA 2793
3254 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC-GGTGATCCCAACAA-TGGAGG 3311
2794 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGGTGATCCCAACAA-TGGAGG 2853
3312 GAAATGTCAGCATGCAATGCGGACCGCTCTGTGCAACACACACAGGCA 3371
2854 GAAATGTCAGCATGCAATGCGGACCGCTCTGTGCAACACACACAGGCA 2913
3372 GTGCTTCTGCACCAACAAAGGGCGTCAAGGGGACAGTCCAGCTATGTGAGGTAGAAAA 3431
2914 GTGCTTCTGCACCAACAAAGGGCGTCAAGGGGACAGTCCAGCTATGTGAGGTAGAAAA 2973
3432 TCGATACAGGAAACCTCTCAGAGGAACATGTTATTTATTTACTCTTCTTATGACTATCA 3491
2974 TCGATACAGGAAACCTCTCAGAGGAACATGTTATTTATTTACTCTTCTTATGACTATCA 3033
3492 GTTCACCTTTTGTCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGC 3551
3034 GTTCACCTTTTGTCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGC 3093
3552 TACTCTGACGAACAAAAACAGGATTTGGAATGTTATCAATGCTCCAGAAATTTCAA 3611
3094 TACTCTGACGAACAAAAACAGGATTTGGAATGTTATCAATGCTCCAGAAATTTCAA 3153
3612 CCTCAACATCAGCTGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCC 3671
3154 CCTCAACATCAGCTGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCC 3213
3672 TGTGTTTCAAAAAACCAACATTAAGGAGTCAAAAGATAGTTTCTCTAATGAGAAGTTGA 3731
3214 TGTGTTTCAAAAAACCAACATTAAGGAGTCAAAAGATAGTTTCTCTAATGAGAAGTTGA 3273
3732 TTTTTCGACCAACCCCAATATCAGTTTCTTTGTTTATGTCAGTAATTTTCACTGGCCAT 3791
3274 TTTTTCGACCAACCCCAATATCAGTTTCTTTGTTTATGTCAGTAATTTTCACTGGCCAT 3333
3792 CAAAATTCAGAT 3803
3334 CAAAATTCAGAT 3345

RESULT 13

AAZ91922

ID AAZ91922 standard; cDNA; 2625 BP.

XX AC AAZ91922;

XX AC AAZ91922;

DT 08-JUN-2000 (first entry)

XX Human mahogany protein coding sequence #4.

XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
XX weight regulation; cell therapy; body weight disorder; cachexia;
XX anorexia; hyperpigmentation; increased metabolic rate disorder;


```
QY 1443 TTTGGATTAAGAACACATCGAGTATATATACACCCAGGGTCCCTTGTGCAAGGGGTTA 1502
DB 1440 TTTGGATTAAGAACACATCGAGTATATATACACCCAGGGTCCCTTGTGCAAGGGGTTA 1499
QY 1503 CGGCCATAGCAGTGTATACGACCATAGGACAGGGCCCTATACGTTTCATGTGTGCTCAAA 1562
DB 1500 CGGCCATAGCAGTGTATACGACCATAGGACAGGGCCCTATACGTTTCATGTGTGCTCAAA 1559
QY 1563 GCGTTTCAGTCCCAATAGTACCGGTTGACAGATCATCTTACCGATATGATGCGATAC 1622
DB 1560 GCGTTTCAGTCCCAATAGTACCGGTTGACAGATCATCTTACCGATATGATGCGATAC 1619
QY 1623 CCAGATGTGGACCATCTTAAGGACAGCCGATTTTCGGTTACTTGCACACAGCTGTGAT 1682
DB 1620 CCAGATGTGGACCATCTTAAGGACAGCCGATTTTCGGTTACTTGCACACAGCTGTGAT 1679
QY 1683 AGTCAGTGGAAACCATGCTGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCA 1742
DB 1680 AGTCAGTGGAAACCATGCTGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCA 1739
QY 1743 TGGCGCCAAATCTCTCTTCAGATTTTCAGTCCCTATGACATTCCTGCTGACCGCTGGTC 1802
DB 1740 TGGCGCCAAATCTCTCTTCAGATTTTCAGTCCCTATGACATTCCTGCTGACCGCTGGTC 1799
QY 1803 AGTCCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTTGGCCATTCAGCAGTCTT 1862
DB 1800 AGTCCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTTGGCCATTCAGCAGTCTT 1859
QY 1863 ACACAAAGCAACCATGATGTTGCGGTGTTCAATAGTCTCTCTCAGCGACATPCCT 1922
DB 1860 ACACAAAGCAACCATGATGTTGCGGTGTTCAATAGTCTCTCTCAGCGACATPCCT 1919
QY 1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTGTTTACGACGAG 1982
DB 1920 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTGTTTACGACGAG 1979
QY 1983 ACCTGGTATTCGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGGCGCTGGC 2042
DB 1980 ACCTGGTATTCGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGGCGCTGGC 2039
QY 2043 AACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCA 2102
DB 2040 AACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCA 2099
QY 2103 TCACAGATGTACACGACACAGATGTTTACAGCTGTACAGCCCAACCAATGACTGCCA 2162
DB 2100 TCACAGATGTACACGACACAGATGTTTACAGCTGTACAGCCCAACCAATGACTGCCA 2159
QY 2163 CTGTTGCAATGACATTTGTGTCGCCAGAACACACAGTGTCTAGAAGCCAGATCTCCAT 2222
DB 2160 CTGTTGCAATGACATTTGTGTCGCCAGAACACACAGTGTCTAGAAGCCAGATCTCCAT 2219
QY 2223 TTTTAGGTATGAGATTTGCCCCAGGATAACCCCATGCTACTGTATACAGAGAGACCAG 2282
DB 2220 TTTTAGGTATGAGATTTGCCCCAGGATAACCCCATGCTACTGTATACAGAGAGACCAG 2279
QY 2283 CTGAGGAGCTGTCCCTTGGACACAGAACTGTCAGTGGAGCCCGGGAATCAGGAGTGCAT 2342
DB 2280 CTGAGGAGCTGTCCCTTGGACACAGAACTGTCAGTGGAGCCCGGGAATCAGGAGTGCAT 2339
QY 2343 TGCCTTCCCGGAAATATCTGTGSCATTTGGTGCATTTGGTTGG 2387
DB 2340 TGCCTTCCCGGTAAGGCTTGCAGGGTCAATCTTGGTGTGTGGG 2384
```

RESULT 14

AAZ91919

ID AAZ91919 standard; cDNA; 6373 BP.

XX AAZ91919;

AC AAZ91919;

XX AAZ91919;

DT 08-JUN-2000 (first entry)

```
XX DE Human mahogany protein coding sequence #1.
XX KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antioesity; antianorexic; anticachexic; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT misc_difference 448 /tag= a
FT /note= "represented in specification as : "
FT misc_difference 615 /tag= b
FT /note= "represented in specification as : "
FT misc_difference 683 /tag= c
FT /note= "represented in specification as : "
XX PN WO200005373-A2.
XX PD 03-FEB-2000.
XX XX 21-JUL-1999; 99WO-US016484.
XX PR 21-JUL-1998; 98US-0093630P.
XX PR 20-OCT-1998; 98US-0104978P.
XX PR 05-FEB-1999; 99US-00245041.
XX PA (MILL-) MILLENIUM PHARM INC.
XX PI Moore K, Nagle DL;
XX XX WPI; 2000-195103/17.
XX DR P-PSDB; AAY81806.
XX PT New human and murine mahogany genes, useful, e.g. for diagnosis and
XX treatment of body weight disorders.
XX PS Claim 1; Fig 10a; 188pp; English.
XX CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (ii) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (ii)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX SQ Sequence 6373 BP; 1736 A; 1468 C; 1460 G; 1699 T; 0 U; 10 Other;
Query Match 42.7%; Score 1830.2; DB 3; Length 6373;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1986; Conservative 7; Mismatches 41; Indels 26; Gaps 11;
QY 2257 ATGTACTACTGTACACAGACAGCAGCTGCGAGGAGCTGCGCTGGACCACTGCCAG 2316
DB 1 ATGTACTACTGTACACAGACAGCAGCTGCGAGGAGCTGCGCTGGACCACTGCCAG 60
QY 2317 TGGGAGCCCCCGGAATCAGAGTGTGCTTCCCTGCCGAAATATCTGTGCACTGGCTGG 2376
```

Db 61 TGGAGCCCCGATCAGAGTGCATTGCCCTGCCGAAATATCTGTGCAATGGCTGG 120
Qy 2377 CATTTGGTTGAAACTCATGTTTGAATAATCTACTGCCAAGAGAAATATGCAATGCT 2436
Db 121 CATTTGGTTGAAACTCATGTTTGAATAATCTACTGCCAAGAGAAATATGCAATGCT 180
Qy 2437 AAATTTGTTCTGTAGAACCAATGCCCTTTGGCTTCTCTTACAACCCGAGAGAGAGTA 2496
Db 181 AAATTTGTTCTGTAGAACCAATGCCCTTTGGCTTCTCTTACAACCCGAGAGAGTA 240
Qy 2497 GAATTTGTTCTTAAGACGCTGCGAATAATGAGTCATCTCAGAGCATGTCCAAAGCTCAC 2556
Db 241 GAATTTGTTCTTAAGACGCTGCGAATAATGAGTCATCTCAGAGCATGTCCAAAGCTCAC 300
Qy 2557 TTAACCCCATGGTTCGGCTT--CGGAGATCAATGTCTTACTGG--TGCTGGG 2608
Db 301 TTAACCCCATGGTTCGGCTT--CGGAGAGGTTCATGTCTTACTTGGTCTGGGGAAG 360
Qy 2609 AAGATATGTCCCCATTTACAATAG-TTTACTACAGTGGATGCC----GTCTGAGCCGAG 2715
Db 361 GATATGTCCTCCATTTTACAATAGTTTACTACAGTGGATGCCGTCTTGAGGCCGAG 420
Qy 2664 TGATGCT--GAATTCGTGGAAATTTAT----CAGAACCCAGT--ACTCGGGAGCTGAAG 2715
Db 421 TGTTCCTTGGAAATCTGTGGAAATTTTNAATTCAGGAACCCAGTTACTTTCGGGAGCTGAAG 480
Qy 2716 GCTGCAACCTGCAT--CAACCCACTCAATGTAGTGTCTGTGAAGGCCCTGCACCAACACA 2773
Db 481 GCTGCAACCTGCATTCACCACTYMAATGGTAGTGTCTGTGAAGGCCCTGCACCAACACA 540
Qy 2774 GTGTAA--GCAGTCCGGACACCATGTGCCCTTGAGGACAGCATGTGAGANTGACCCAGC 2832
Db 541 GTGTAAAGGAGTCCGGACACCATGTGCCCTTGAGGACAGCATGTGAGANTGACCCAGC 600
Qy 2833 GGCAGCTGTAGTG-CATGTGGTGCAGCAATCAAGCAGTGTTGGACTCCCATGCTTA 2891
Db 601 GGCAGCTGTAGTGNCATGTGGTGCAGCAATCAAGCAGTGTTGGACTCCCATGCTTA 660
Qy 2892 TGTGGCTCTTCCCTTTTGG-CCAGTGTATGGAATGGTATACGATGAGCACCTTGGCCCC 2950
Db 661 TGTGGCTCTTCCCTTTTGGNCCAGTGTATGGAATGGTATACGATGAGCACCTTGGCCCC 720
Qy 2951 CTGAATAATGTTGAGCTTACTGTACTGTACTGTTTGGAGCAACAGGCTGTGGCT 3010
Db 721 CTGAATAATGTTGAGGCTTACTGTACTGTACTGTTTGGAGCAACAGGCTGTGGCT 780
Qy 3011 GGTGTACTGTATCCAGCAATCTGCAAGGAAATGCAATAGAGGGTTCTTATAAAGGAC 3070
Db 781 GGTGTACTGTATCCAGCAATCTGCAAGGAAATGCAATAGAGGGTTCTTATAAAGGAC 840
Qy 3071 CAGTGAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATFCCAAGCCCTGTCTCAATT 3130
Db 841 CAGTGAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATFCCAAGCCCTGTCTCAATT 900
Qy 3131 CCAGCATGTCTAGAGCAGCAGATACACTGCTTTCATTCATCTGCTCCAGCTTGGC 3190
Db 901 CCAGCATGTCTAGAGCAGCAGATACACTGCTTTCATTCATCTGCTCCAGCTTGGC 960
Qy 3191 AATGCAACGGCCACAGTAAATGCATCAATCAGACATCTGTGAGAAGTGTGAGAACCTGA 3250
Db 961 AATGCAACGGCCACAGTAAATGCATCAATCAGACATCTGTGAGAAGTGTGAGAACCTGA 1020
Qy 3251 CCACGGAAGCAGCTCCAGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAG 3310
Db 1021 CCACGGAAGCAGCTCCAGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAG 1080
Qy 3311 GGAATGTGAGCATGCAAGTGCATAGGCGACGCTCTCTGTGCAACCAACCAACGGGCA 3370
Db 1081 GGAATGTGAGCATGCAAGTGCATAGGCGACGCTCTCTGTGCAACCAACCAACGGGCA 1140
Qy 3371 AGTGTCTGTGCAACCAAGGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAA 3430

Db 1141 AGTGTCTTGTGCAACCAAGGGCGTCAAGGGGGACGAGTGCAGCTATGTGAGGTAGAAA 1200
Qy 3431 ATCGATACCAAGAAACCTCTCAGAGGAACATGTTATATATCTCTTCTTATGACTATC 3490
Db 1201 ATCGATACCAAGAAACCTCTCAGAGGAACATGTTATATATCTCTTCTTATGACTATC 1260
Qy 3491 AGTTCACCTTTAGTCTATCCCAAGGAAGATGATCGCTATTACACAGCTATCAATTTTGG 3550
Db 1261 AGTTCACCTTTAGTCTATCCCAAGGAAGATGATCGCTATTACACAGCTATCAATTTTGG 1320
Qy 3551 CTACTCTGACGAACAAACAGGGAATTTGGACATGTTCAATCAATGCTTCCCAAGATTTCA 3610
Db 1321 CTACTCTGACGAACAAACAGGGAATTTGGACATGTTCAATCAATGCTTCCCAAGATTTCA 1380
Qy 3611 ACCTCAACATCACTTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGC 3670
Db 1381 ACCTCAACATCACTTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGC 1440
Qy 3671 CTGTTGTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3730
Db 1441 CTGTTGTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 1500
Qy 3731 ATTTTGGCAACCAACCAATATCACTTCTTGTGTTTATGTCAGTAAATTTCACTGGCCCA 3790
Db 1501 ATTTTGGCAACCAACCAATATCACTTCTTGTGTTTATGTCAGTAAATTTCACTGGCCCA 1560
Qy 3791 TCAAAATTCAGATTCCTTCTCTCAGCAGACAAATTTTATGGAACCTGTGTACAGTTCTTG 3850
Db 1561 TCAAAATTCAGATTCCTTCTCTCAGCAGACAAATTTTATGGAACCTGTGTACAGTTCTTG 1620
Qy 3851 TGACTTCTCTAGTTGTTTCTCTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3910
Db 1621 TGACTTCTCTAGTTGTTTCTCTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Qy 3911 AACAAAGTTTGGGGCTCCAGACGTAGAGAGCAACTTCTTCCAGAGATGCAACAGATGG 3970
Db 1681 AACAAAGTTTGGGGCTCCAGACGTAGAGAGCAACTTCTTCCAGAGATGCAACAGATGG 1740
Qy 3971 CGAGCGTCCCTTGTGCTCTGTAAATGTCGCTTGGAAACAGATGAGGAGCCTCTGATC 4030
Db 1741 CGAGCGTCCCTTGTGCTCTGTAAATGTCGCTTGGAAACAGATGAGGAGCCTCTGATC 1800
Qy 4031 TTATTTGGGGGAGTATAAAGACTGTTTCCCAACCAATGCACTGGAGCGCTGTTTGGCA 4090
Db 1801 TTATTTGGGGGAGTATAAAGACTGTTTCCCAACCAATGCACTGGAGCGCTGTTTGGCA 1860
Qy 4091 ACAAGCCCTGTCTCTCTGTGTTTGTGAGGCTCCCTGAGGCTGGGTGGATCCCTC 4150
Db 1861 ACAAGCCCTGTCTCTCTGTGTTTGTGAGGCTCCCTGAGGCTGGGTGGATCCCTC 1920
Qy 4151 CTCTTGGGCACTCAGGTCTTGTGCTGGCCAGCGCTGCTGAGCATTTCTCAGCAGATGC 4210
Db 1921 CTCTTGGGCACTCAGGTCTTGTGCTGGCCAGCGCTGCTGAGCATTTCTCAGCAGATGC 1980
Qy 4211 CGATAGTGTACAAAGGAGAGTCAAGGCGCTGAGAAACCGGAAGCAGCAGCCCCCTGAC 4270
Db 1981 CGATAGTGTACAAAGGAGAGTCAAGGCGCTGAGAAACCGGAAGCAGCAGCCCCCTGAC 2040
Qy 4271 AGCTTGGGACCTGCATCTGA 4290
Db 2041 AGCTTGGGACCTGCATCTGA 2060

RESULT 15

AAZ91917

ID AAZ91917 standard; cDNA; 2419 BP.

XX

AAZ91917;

XX

08-JUN-2000 (first entry)

XX

Murine mahogany protein coding sequence akml003.

XX

KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Artichobesity; antianorexic; anticachexic; ss.
XX
OS Mus sp.
XX
PN WC200005373-A2.
XX
XX
PD 03-FEB-2000.
XX
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
XX Moore K, Nagle DL;
XX
XX WPI; 2000-195103/17.
DR P-PSDB; AAY81804.
XX
XX New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
XX
PS Claim 1; Fig 8a; 189pp; English.
XX
XX This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iv) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (ii) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (ii) for diagnosis, prognosis or monitoring of treatment; to
CC evaluate (ii)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;
Query Match 35.5%; Score 1521.8; DB 3; Length 2419;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 1895; Conservative 0; Mismatches 247; Indels 205; Gaps 2;
3 GGTGGCGCGCGCGCGCACTGAGGCAAGGCTCAGGAGGAGGACCGCGCGCGGCGGCGC 62
82 GGTGGCGCGCGCGCGCGCACTGAGGCGCGGCTCAGGCGGAGCGACGAGGAGCAGCAGC 141
63 GCTCGCGGCGAGGACGCGCGCGCGCGCACTGAGGAGCTGGGAGCTGGGAGCGCTGGAGGCC 122
142 GCCTCGCGGCGAGGAGGC 201
123 GGGGCTGGGGGCGCGCGCTGCGCGCTCCGCGCGCTGCTCTCCACCGCTGGCGGCGCGCGC 182
202 GGGACCGCGCGCGCGCTGCTCTCCGCGGCTGCTCTCGCGGCGCTGCGCGCGCGCGC 261
183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
243 CGAGCG 302
313 CGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372

QY 303 TGACCGGCGCTGTGTCAACGGCGGTGCTGTCAACCCCTGGCACCGCCAGTGGCTTGCC 362
DB 373 TGACCGGCGCTGTGTCAACGGCGGTGCTGTCAACCCCTGGCACCGCCAGTGGCTTGCC 432
QY 363 CCGCGGCTGGGTGGGCGAGCAATGCCAGCATCTGCGGGGCGCGCTTCAG- 410
DB 433 CACGGGCTGGGTGGGCGAGCAATGCCAGCATCTGCGGGGCGCGCTTCAGGACATCTGCTC 492
QY 411 ----- 410
DB 493 AGCGCTATAATCACAGCTGTTCCGGAAGGTGAGGCTGGAGGAACAGTTGCGAGCAAGCTTC 552
QY 411 ----- 410
DB 553 GGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTCCAAACCAAAA 612
QY 411 ----- 410
DB 613 TGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTGACTAATTTTTCACC 672
QY 411 ----ACTRACTGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAATAACAAA 466
DB 673 GAGAATACTGCTCTTCTGGAATTTGTAAAGATGGACCTGGGAAATTTATAATAAGA 732
QY 467 CGAAGTGACGCTGCTCATTTAGGACAGCGCAATAGAAATAGATGAGCTTCTGTTCAATC 526
DB 733 CGAAGTGACATGGCTCATTTAGGACAGCGCAATAGAAATAGATGAGCTTCTGCTTCAACC 792
QY 527 ATTTTGTACACAGTGTAGTTGGGACCAATTTATGTTTATGATGGGACCTCAATTTATG 586
DB 793 ATTTTGTACAGAAATGATGCTGGGACCAATTTATGTTTATGATGGGACCTCAATTTACG 852
QY 587 CACCGCTAGTCTGCTCATTTAGTGGCTCATTTGTTTCTCTGAGAGAGATCGCAATGAGCTG 646
DB 853 CACCTCTGATCTGCTGCTTTAGTGGCTCATTTGTTTCTCTGAGAGAGATCGCAATGAGCTG 912
QY 647 TCCTTGAGGTTGTGCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
DB 913 CTCTTGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
QY 707 ATAAATTTGACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 766
DB 973 ATAAATTTGACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1032
QY 767 GCGGAGGAGTGTAAAGTCAATGATGAGGAGGAACTGTTGAATGTGAATGTGTTGAA 826
DB 1033 GCGGAGGAGTGTAAAGTCAATGATGAGGAGGAACTGTTGAATGTGAATGTGTTGAA 1092
QY 827 ACTGGAAGGTGAAGCATGTGACATTTCTCTACTGTACAGACAATCTGTTGTTTCTCATC 886
DB 1093 ACTGGAAGGTGAAGCATGTGACATTTCTCTACTGTACAGACAATCTGTTGTTTCTCATC 1152
QY 887 GAGGCATCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 946
DB 1153 GAGGCATCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1212
QY 947 CTGGATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1006
DB 1213 CTGGATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1272
QY 1007 ACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1066
DB 1273 ACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1332
QY 1067 TTGGAGGATATATGTTCAACCACTCAGATTAATTAATTAATTAATTAATTAATTAATTA 1126
DB 1333 TTGGAGGATATATGTTCAACCACTCAGATTAATTAATTAATTAATTAATTAATTAATTA 1392
QY 1127 CTAGGAGTGGTTCCTCACTAAACCGTCTGTTGAACAATGTTGTTGTTGTTGTTGTTGTT 1186
DB 1393 CTAGGAGTGGTTCCTCACTAAACCGTCTGTTGAACAATGTTGTTGTTGTTGTTGTTGTT 1452

Qy 1187 CTTTGGCAATTATACAGGATAAATTTACATGTATGGAGGAAAAATGATCCAACTGGGA 1246
Db |||||
Qy 1453 CTTTGGCAATTACATAAGATAAATCTACATGTATGGAGGAAAAATGATTCACAGGGA 1512
Db |||||
Qy 1247 ATGTGACCAATGAGTTGAGAGTTTTTTCACATTCATAATGAGTCATGGGTGTTGTGACCC 1306
Db |||||
Qy 1513 ACGTGACCAATGAGCTGAGAGTATTTTCATATTCATATGATCATGGGTATGTTAACTC 1572
Db |||||
Qy 1307 CTAAGGCAAGAGGAGCAGTATGACGTGTTGGGCACTCTGCACACATTTGACATGAAGA 1366
Db |||||
Qy 1573 CGAAAGCTAAGGATCAGTATGACGTGTTGGACACTCAGCACACATTTTACACTGGCAT 1632
Db |||||
Qy 1367 ATGGCCGAGTGTCTATGCTGTCATCTTGTGTCATCTCTCTCTATGATATATAAGCA 1426
Db |||||
Qy 1633 CTGGCCGTGTGGTCATGTTGGTCACTCTCGTCACTGCTGCTATGATATATAAGCG 1692
Db |||||
Qy 1427 ATGTGACGGAATATGATTTGGATAAGAACACATGGAGTATATACACACCCAGGGTGCC 1486
Db |||||
Qy 1693 TTGTGCAAGGAATATGACTTGGAAAAAGAACACATGGAGTATATACATACCTCAGGGTGCTC 1752
Db |||||
Qy 1487 TTGTGCAAGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCCAGGGCCCTATACG 1546
Db |||||
Qy 1753 TTGTGCAAGGGGTTATGGCCACAGTAGTGTATGATGACAGGACCAAGGCTCTGTACG 1812
Db |||||
Qy 1547 TTCAATGTTGGCTACAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGACAGTATCTCTACC 1606
Db |||||
Qy 1813 TTCAATGTTGGCTACAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGACAGTATCTCTACA 1872
Db |||||
Qy 1607 GATATGATGTGGATACCCAGATGTGGACCATCTTAAAGACAGCCGATTTTTCGGTTACT 1666
Db |||||
Qy 1873 GATACGATGTGGATACCTAGTGTGGACCATCTTAAAGACAGCCGATTTTTCGGTTACT 1932
Db |||||
Qy 1667 TGCAACAGCTGTGATAGTGTGAGTGGAAACCACTGCTGTTGTTGGGGGAAACACACACAATG 1726
Db |||||
Qy 1933 TGCAATACGCTGTGATAGTGTGAGTGGAAACCACTGCTGTTGTTGGGGGAAACACACACAATG 1992
Db |||||
Qy 1727 ACACATCTATGAGCCATGGGCCAAATGCTTCTCTTCCAGATTTTCATGGCCCTATGACATTG 1786
Db |||||
Qy 1993 ACACCTTCCATGAGCCACGGTGCCAAATGCTTCTCTCTGAGCTTTCATGGCTTATGACATTG 2052
Db |||||
Qy 1787 CCTGTGACCGTGTGCTAGTGTCTCCAGACCTGATCTCCACCATGATGTCACACAGATTG 1846
Db |||||
Qy 2053 CTTGTGACCGATGTGCTAGTGTCTCCAGACCTGATCTCCATCATGATGTCAACAGATTG 2112
Db |||||
Qy 1847 GCCATTGACAGTCTTTACACACAGCACCATGTAATGTTGCGGTGTTTCAATAGTCTCC 1906
Db |||||
Qy 2113 GCCATTGACAGTCTTTGTAACACAGCACCATGTAATGTTGCGGTGTTTCAACAGCCTCC 2172
Db |||||
Qy 1907 TCCTCAGGACATCTGTGTTATTCCTCGGACAGTGTGATGCGCATCGGAGTGAAGCCG 1966
Db |||||
Qy 2173 TCCTCAGTGAAGTCTTGTGTTTACCTCGGAGCAGTGGATGACACCCGAGTGAAGCTG 2232
Db |||||
Qy 1967 CTTGTTTACGACGAGGACCTGTTATTCGTTGTGTGTAACACAGGTCGTTCTCAGTGTA 2026
Db |||||
Qy 2233 CTTGTTGAGCAGGAGGACCTGTTATTCGTTGTGTGTGGACACACAGTCTCTCGATGTA 2292
Db |||||
Qy 2027 TCTCGTGGCGCTGGCACTGATGAACAGAGGAAAAATGTTAAATCAGATGTTTTTCCA 2086
Db |||||
Qy 2293 CCTCTGGGAGTGTGCAACTGAAGAACAAAGCAGAAAAATGTTAAATCAGAGTGTTTTTCTA 2352
Db |||||
Qy 2087 AAGGACCTTTGACCATGACAGATGTGACAGCAGCAGATTTGTTACAGCTGTACGCCA 2146
Db |||||
Qy 2353 AAGGACCCCTTGACCATGACAGATGTGACAGCAGCAGATTTGTTACAGCTGTACGCCA 2412
Db |||||
Qy 2147 ACACCAA 2153
Db |||||
Qy 2413 ATACCAA 2419
Db |||||

Result No.	Score	Query Match	Length	DB	ID	Description
1	4184.8	97.5	8589	9	US-09-893-238-14	Sequence 1
2	3696.2	86.2	4072	9	US-09-893-238-16	Sequence 1
3	3398	79.2	8827	9	US-09-893-238-1	Sequence 1
4	2248.2	52.4	2625	9	US-09-893-238-18	Sequence 1
5	1853	43.2	6370	9	US-09-893-238-12	Sequence 1
6	1521.8	35.5	2419	9	US-09-893-238-8	Sequence 8
7	1463.8	34.1	6733	16	US-10-197-824-1	Sequence 1
8	640.2	14.9	1051	9	US-09-893-238-10	Sequence 1
C 9	277.4	6.5	625	14	US-10-198-846-5869	Sequence 9
C 10	249.2	5.8	531	9	US-09-864-761-15900	Sequence 1
C 11	246	5.7	246	9	US-09-736-457-966	Sequence 9
C 12	246	5.7	246	9	US-09-502-541-966	Sequence 9
C 13	246	5.7	246	9	US-09-849-622-966	Sequence 9
C 14	246	5.7	246	14	US-10-017-754-966	Sequence 9
C 15	246	5.7	246	14	US-10-113-872-966	Sequence 9

Db 2280 CTGCAGAGCTGTGCCCTCGACAGAACTCCAGTGGGAGCCCGGNAATCAGAGTGCAT 2339
Qy 2343 TGGCTCCGCCAAATAATCTGTGGCAATGGCTGGCAATTTGGTTGGAAATCATGTTTGAA 2402
Db 2340 TGGCTCCGCCAAATAATCTGTGGCAATGGCTGGCAATTTGGTTGGAAATCATGTTTGAA 2399
Qy 2403 AATTACTACTGCCAAGAGAAATATGCAATGCTAAATTTCTGTGTAGGAACCAACATGC 2462
Db 2400 AATTACTACTGCCAAGAGAAATATGCAATGCTAAATTTCTGTGTAGGAACCAACATGC 2459
Qy 2463 CCTTTTGGCTTCTCTTACAAACCCAGAGAAAGTAGAATTTGTCTTTAAGCAGCTGCGAAT 2522
Db 2460 CCTTTTGGCTTCTCTTACAAACCCAGAGAAAGTAGAATTTGTCTTTAAGCAGCTGCGAAT 2519
Qy 2523 AATGCAATCTCTCAGAGCAATGTCAGCTCACTTAACCCCAATGGTGGCCCTTCGGA 2582
Db 2520 AATGCAATCTCTCAGAGCAATGTCAGCTCACTTAACCCCAATGGTGGCCCTTCGGA 2579
Qy 2583 GATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCCAATTTACAAATAGTTTACTACA 2642
Db 2580 GATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCCAATTTACAAATAGTTTACTACA 2639
Qy 2643 GTGATCCGCTGTGAGCCCAAGTAGATGCTGGAATCTGTGGAATTTTATCAGAACCCAGTAC 2702
Db 2640 GTGATCCGCTGTGAGCCCAAGTAGATGCTGGAATTTTATCAGAACCCAGTAC 2699
Qy 2703 TCGGGACTGAGGCTTGCACCTGCTCAACCCCAATGCTGTGTGAAAGGCC 2762
Db 2700 TCGGGACTGAGGCTTGCACCTGCTCAACCCCAATGCTGTGTGAAAGGCC 2759
Qy 2763 TGCAAAACCAAGTCTAAGCAGTSCCGGACACCATGTGCTTGAAGCAGCATGTGAGAA 2822
Db 2760 TGCAAAACCAAGTCTAAGCAGTSCCGGACACCATGTGCTTGAAGCAGCATGTGAGAA 2819
Qy 2823 TTGCACAGCGGAGCTCTGAGTGCATGTGTCAGCAACCATGAGCAGTGTGAGACTC 2882
Db 2820 TTGCACAGCGGAGCTCTGAGTGCATGTGTCAGCAACCATGAGCAGTGTGAGACTC 2879
Qy 2883 CAATGCTCTATGTGGCTCTCTTCCCTTTGGCCAGTATGGAATGATATGATGAGCAGC 2942
Db 2880 CAATGCTCTATGTGGCTCTCTTCCCTTTGGCCAGTATGGAATGATATGATGAGCAGC 2939
Qy 2943 CTGCCCCCTGAAATTTGTCAGCTACTGTA CTTGTAGTCAATGCTTGGAGCAACAGG 3002
Db 2940 CTGCCCCCTGAAATTTGTCAGCTACTGTA CTTGTAGTCAATGCTTGGAGCAACAGG 2999
Qy 3003 CTGTGGCTGTGTACTGATCCCAATATCTGGCAAGGAAATGATAGAGGTTCCCTA 3062
Db 3000 CTGTGGCTGTGTACTGATCCCAATATCTGGCAAGGAAATGATAGAGGTTCCCTA 3059
Qy 3063 TAAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122
Db 3060 TAAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119
Qy 3123 GCTCAATTCAGCATGTGCTTAGAGGACAGAGATCAACTGTGCTTTTCAATTCATGCTCC 3182
Db 3120 GCTCAATTCAGCATGTGCTTAGAGGACAGAGATCAACTGTGCTTTTCAATTCATGCTCC 3179
Qy 3183 AGCTTGCAATGCAAGCCCAAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGA 3242
Db 3180 AGCTTGCAATGCAAGCCCAAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGA 3239
Qy 3243 GAACTGACCAAGGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTATCCAC 3302
Db 3240 GAACTGACCAAGGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTATCCAC 3299
Qy 3303 CAATGGAGGAAATGTCAGCCATCAAGTGAATGGGCAAGCCGCTCTGTGCAACCAAA 3362
Db 3300 CAATGGAGGAAATGTCAGCCATCAAGTGAATGGGCAAGCCGCTCTGTGCAACCAAA 3359
Qy 3363 CACGGCAAGTGTCTTGCACCAACAAAGGCGCTCAAGGGGACAGTGCAGCTATGTA 3422
Db 3360 CACGGCAAGTGTCTTGCACCAACAAAGGCGCTCAAGGGGACAGTGCAGCTATGTA 3419

RESULT 2

US-09-893-238-16

; Sequence 16, Application US/09893238

; Patent No. US20020150973A1

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; FILE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

; FILE REFERENCE: 7853-237

; CURRENT APPLICATION NUMBER: US/09/893,238

Qy 3423 GGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3482
Db 3420 GGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3479
Qy 3483 TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAA 3542
Db 3480 TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAA 3539
Qy 3543 TTTTGTGGCTACTCTGAGCAACAAACAGGATTTGACATGTTCAATGCTGCTCAA 3602
Db 3540 TTTTGTGGCTACTCTGAGCAACAAACAGGATTTGACATGTTCAATGCTGCTCAA 3599
Qy 3603 GAATTTCAACCTCAACATCACTGCGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662
Db 3600 GAATTTCAACCTCAACATCACTGCGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3659
Qy 3663 AGAGATGCTGTTTCAAAAACCAACATTAAGAGGATACAAAGATAGTTTCTCTAATGA 3722
Db 3660 AGAGATGCTGTTTCAAAAACCAACATTAAGAGGATACAAAGATAGTTTCTCTAATGA 3719
Qy 3723 GAAGTTTGAATTTCCGAACCCCAATATCACTTTCTTTGTTATGTGAGTAATTTCCAC 3782
Db 3720 GAAGTTTGAATTTCCGAACCCCAATATCACTTTCTTTGTTATGTGAGTAATTTCCAC 3779
Qy 3783 CTGGCCCATCAAAATTCAGATTGCCCTTCTCAGCACAGCAATTTTATGGACCTGGTACA 3842
Db 3780 CTGGCCCATCAAAATTCAGATTGCCCTTCTCAGCACAGCAATTTTATGGACCTGGTACA 3839
Qy 3843 GTTCTTGTGACTTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3902
Db 3840 GTTCTTGTGACTTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3899
Qy 3903 GAAGATCAAAACAAAGTTGTTGGCCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCA 3962
Db 3900 GAAGATCAAAACAAAGTTGTTGGCCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCA 3959
Qy 3963 ACAGATGGCCAGCGCTCCCTTTGCCCTCTGTAATGTGCGCTTGGAAAACAGATGAGAGCC 4022
Db 3960 ACAGATGGCCAGCGCTCCCTTTGCCCTCTGTAATGTGCGCTTGGAAAACAGATGAGAGCC 4019
Qy 4023 TCTGATCTTATTTGGGGGAGTATAAGACTTCTCCCAACCCATTCACATGCACTGGAGCCGTG 4082
Db 4020 TCTGATCTTATTTGGGGGAGTATAAGACTTCTCCCAACCCATTCACATGCACTGGAGCCGTG 4079
Qy 4083 TTTTGGCAACAAAGCGCTGCTCTCTGTGTTGTAGGCTCCCTCGAGGCTCTGGGTGG 4142
Db 4080 TTTTGGCAACAAAGCGCTGCTCTCTGTGTTGTAGGCTCCCTCGAGGCTCTGGGTGG 4139
Qy 4143 CATCCCTCTCTCTGGCAGTCAAGTCTTCTGCTGGCCAGCGCCCTGTTGACATTTCTCA 4202
Db 4140 CATCCCTCTCTCTGGCAGTCAAGTCTTCTGCTGGCCAGCGCCCTGTTGACATTTCTCA 4199
Qy 4203 GCAGATGCCGATAGTGTACAAGGAAGTCAAGAGCCCTGAGAAAACCCGGAACAGCAGCC 4262
Db 4200 GCAGATGCCGATAGTGTACAAGGAAGTCAAGAGCCCTGAGAAAACCCGGAACAGCAGCC 4259
Qy 4263 CCTGCACAGCCTGGGACCTGCATCTGA 4290
Db 4260 CCTGCACAGCCTGGGACCTGCATCTGA 4287

; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16

; LENGTH: 4072

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-893-238-16

Query Match 86.2%; Score 3696.2; DB 9; Length 4072;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 3744; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
3 GGTGCGCCGACGCGCGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGAGCGGAGC 62
9 GGTGCGCGCGCGCGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGGAGC 68
63 GCTGCGCGCGCGCGCGCGCGCACTGGGACTGGGAGCTGACAGGAGGCTGGGAGGCC 122
69 GCTGCG 128
123 GGGGCTGGGGCG 182
129 GGGACCG 188
183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
189 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
243 CGAGCG 302
240 CGAGCG 239
303 TGACCG 362
300 TGACCG 359
363 CG 422
360 CG 419
423 TTCTGGGTTTGTGACAGATGGAAGCTGGAATTTAATAATAAAACGAAAGTGCACGTGGCT 482
420 TTCTGGGTTTGTGACAGATGGAAGCTGGAATTTAATAATAAAACGAAAGTGCACGTGGCT 479
483 CATTGAGGACGAGCCCAATAGAAATAGAGATGCTGCTTCAATCAATTTGCTACAGAGTG 542
480 CATTGAGGACGAGCCCAATAGAAATAGAGATGCTGCTTCAATCAATTTGCTACAGAGTG 539
543 TAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGC 602
540 TAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGC 599
603 ATTATGAGGCTCAATTTGCTGAGAGAGATGGGAGATGAGACTGCTGAGAGTTGCTGC 662
600 ATTATGAGGCTCAATTTGCTGAGAGAGATGGGAGATGAGACTGCTGAGAGTTGCTGC 659
663 CACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
660 CACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
723 TAATATTTACTTACAGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
720 TAATATTTACTTACAGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779
783 GATCAGTAATAGCAGCGAAAGCTGTTGAATGTGAATGTTCTGAAAGCTGGAAGGTCAGC 842

780 GATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAGCTGGAAGGTGAAGC 839
843 ATGTGACATTTCTCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGCATCTGCAATTC 902
840 ATGTGACATTTCTCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGCATCTGCAATTC 899
903 AAGTGATGTACAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTTCC 962
900 AAGTGATGTACAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTTCC 959
963 TGTACCCAGTAAACCACTGATTTGAGCTCGAGAGGATTTCTAACTTAAAGCTCCCCAG 1022
960 TGTACCCAGTAAACCACTGATTTGAGCTCGAGAGGATTTCTAACTTAAAGCTCCCCAG 1019
1023 AGCATCTCAATAAGCTGTGGTCAATGGAAACATTTATGTTGGGTTGTTGAGGATATATGTT 1082
1020 AGCATCTCAATAAGCTGTGGTCAATGGAAACATTTATGTTGGGTTGTTGAGGATATATGTT 1079
1083 CACCACTCAGATTAATACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCC 1142
1080 CACCACTCAGATTAATACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCC 1139
1143 ACTAAACCGTTCTGTGAACAATGTTGTTAGATATGTTGTTGATGTTGTTGTTGTTGTTG 1202
1140 ACTAAACCGTTCTGTGAACAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1199
1203 GGATAAATTTACATGTTATGGAGGAAATTTGATCCAACTGGGATGTGACCAATGATTT 1262
1200 GGATAAATTTACATGTTATGGAGGAAATTTGATCCAACTGGGATGTGACCAATGATTT 1259
1263 GAGAGTTTTCACATTTACATGTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1322
1260 GAGAGTTTTCACATTTACATGTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1319
1323 GTATCAGTGTGTGGGCACTCTGCACACATTTGTTTACCTGAAGAATGTCGAGTGGTCA 1382
1320 GTATCAGTGTGTGGGCACTCTGCACACATTTGTTTACCTGAAGAATGTCGAGTGGTCA 1379
1383 GCTGTCATCTTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
1380 GCTGTCATCTTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
1443 TTTGATTAAGAACACATGAGTATATTAACACCCAGGCTGCTGCTGCTGCTGCTGCTGCT 1502
1440 TTTGATTAAGAACACATGAGTATATTAACACCCAGGCTGCTGCTGCTGCTGCTGCTGCT 1499
1503 CGGCCATAGCAGTGTGTTAGGACCATAGGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562
1500 CGGCCATAGCAGTGTGTTAGGACCATAGGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1559
1563 GGCCTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTTGATAC 1622
1560 GGCCTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTTGATAC 1619
1623 CCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTTCGTTTACTTTCGACACAGCTGTGAT 1682
1620 CCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTTCGTTTACTTTCGACACAGCTGTGAT 1679
1683 AGTGAGTGGACCATTTCTTAAGGACAGCCGATTTTTCGTTTACTTTCGACACAGCTGTGAT 1742
1680 AGTGAGTGGACCATTTCTTAAGGACAGCCGATTTTTCGTTTACTTTCGACACAGCTGTGAT 1739
1743 TGGCGCCCAATGCTTCTCTCAGATTTTCATGGCTTATGACATTTGCTGTTGACCTGCTGCT 1802
1740 TGGCGCCCAATGCTTCTCTCAGATTTTCATGGCTTATGACATTTGCTGTTGACCTGCTGCT 1799
1803 AGTGCTTCCAGACCTGATCTCCACCATGATGTTCAACAGATTTGGCCATTTGAGGCTT 1862
1800 AGTGCTTCCAGACCTGATCTCCACCATGATGTTCAACAGATTTGGCCATTTGAGGCTT 1859
1863 ACACAAACAGCAGCATGATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1922
1860 ACACAAACAGCAGCATGATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919

1923 GGTATTCACTCGGAACAGTGTGATCGGCATCGAGTGAAGCGCTGTTTAGCAGCAGG 1982
1920 GGTATTCACTCGGAACAGTGTGATCGGCATCGAGTGAAGCGCTGTTTAGCAGCAGG 1979
1983 ACCTGGTATTGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGC 2042
1980 ACCTGGTATTGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGC 2039
2043 AACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCACAAAGAACTCTTGACCA 2102
2040 AACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCACAAAGAACTCTTGACCA 2099
2103 TGACAGATGTGACAGCAGACACAGATTGTTACAGCTGTACAGGCCAACACCAATGACCTGCCA 2162
2100 TGACAGATGTGACAGCAGACACAGATTGTTACAGCTGTACAGGCCAACACCAATGACCTGCCA 2159
2163 CTGGTGAATGACCAATGTGTCCAGGAACACAGCTGTCTCAGAGGCCAGATCTCCAT 2222
2160 CTGGTGAATGACCAATGTGTCCAGGAACACAGCTGTCTCAGAGGCCAGATCTCCAT 2219
2223 TTTTAGTATGAGAAATGCCCCAAGGATACCCATGCTACTCTGTAAACAGAGACCCAG 2282
2220 TTTTAGTATGAGAAATGCCCCAAGGATACCCATGCTACTCTGTAAACAGAGACCCAG 2279
2283 CTGACGAGCTGTGCCCTGGACACAGAACTGCGAGTGGAGCCCGGGAATCAGGAGTGCAT 2342
2280 CTGACGAGCTGTGCCCTGGACACAGAACTGCGAGTGGAGCCCGGGAATCAGGAGTGCAT 2339
2343 TGCCTCGCCGAAATATCTGTGGCAATGCGTGGCAATTTGGTGGAACTCATGTTTGA 2402
2340 TGCCTCGCCGAAATATCTGTGGCAATGCGTGGCAATTTGGTGGAACTCATGTTTGA 2399
2403 AATTACTACTGCCAAGAGAAATATGCAATGCTTAAATGTTCTGTAGGAACACCAATGC 2462
2400 AATTACTACTGCCAAGAGAAATATGCAATGCTTAAATGTTCTGTAGGAACACCAATGC 2459
2463 CCTTTGGCTTCTCTTACAAACCCAGAGAGAGTGAATTTGCTTAAAGCAGCTGCGAAT 2522
2460 CCTTTGGCTTCTCTTACAAACCCAGAGAGAGTGAATTTGCTTAAAGCAGCTGCGAAT 2519
2523 AATGAGTCACTCAGAGCAGTGTCCAGCTCACCTTAAACCCAGTGGTGGCCCTCGGAA 2582
2520 AATGAGTCACTCAGAGCAGTGTCCAGCTCACCTTAAACCCAGTGGTGGCCCTCGGAA 2579
2583 GATCAATGTGCTTACTGCTGGAGAGATATGCTCCCAATTTACAAATAGTTTACTACA 2642
2580 GATCAATGTGCTTACTGCTGGAGAGATATGCTCCCAATTTACAAATAGTTTACTACA 2639
2643 GTGATGCGCTCAGGCCAGTGTGAGTCTGAGTCTGTGGAATTTTATCAGAACCCAGTAC 2702
2640 GTGATGCGCTCAGGCCAGTGTGAGTCTGAGTCTGTGGAATTTTATCAGAACCCAGTAC 2699
2703 TCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCC 2762
2700 TCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCC 2759
2763 TGCACCAACAGTGTAAAGAGTGGCCGACACCAATGTCCTTGGAGCAGCATGTGAGA 2822
2760 TGCACCAACAGTGTAAAGAGTGGCCGACACCAATGTCCTTGGAGCAGCATGTGAGA 2819
2823 TTGGACAGCGCAGCTCTGAGTGCATCTGTGTGAGCAACATGAAGCAGTGTGTGGACTC 2882
2820 TTGGACAGCGCAGCTCTGAGTGCATCTGTGTGAGCAACATGAAGCAGTGTGTGGACTC 2879
2883 CAATGCCATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2942
2880 CAATGCCATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2939
2943 CTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCCAGG 3002
2940 CTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCCAGG 2999

3003 CTGTGGCTGGTGTACTGATCCAGCAATPACTGCGAAAGGAAATGCAATAGAGGTTCTTA 3062
3000 CTGTGGCTGGTGTACTGATCCAGCAATPACTGCGAAAGGAAATGCAATAGAGGTTCTTA 3059
3063 TAAAGGACACAGTGAAGTGCCTTGCAGGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122
3060 TAAAGGACACAGTGAAGTGCCTTGCAGGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119
3123 GCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATTCAGTGC 3182
3120 GCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATTCAGTGC 3179
3183 AGCTTGGCAATGCAACAGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAAGTGTGA 3242
3180 AGCTTGGCAATGCAACAGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAAGTGTGA 3239
3243 GAACTGTGCCACAGCAGAGCAGTGCAGACCTGCGATATCTGGCTTCTACGCTATCCAC 3302
3240 GAACTGTGCCACAGCAGAGCAGTGCAGACCTGCGATATCTGGCTTCTACGCTATCCAC 3299
3303 CAATGGAGGAAATGTCAGCCATGCAAGTGAATGGGCAACGCGTCTGTGTGCAACACCAA 3362
3300 CAATGGAGGAAATGTCAGCCATGCAAGTGAATGGGCAACGCGTCTGTGTGCAACACCAA 3359
3363 CACGGGCAAGTCTTCTGCACCCACCAAGGGGTGCAAGGGGACAGTGCAGCTATGTA 3422
3360 CACGGGCAAGTCTTCTGCACCCACCAAGGGGTGCAAGGGGACAGTGCAGCTATGTA 3419
3423 GGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTACTCTTTCTTAT 3482
3420 GGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTACTCTTTCTTAT 3479
3483 TGACTATCAGTTCACTTGTAGTCTATCCAGGAAATGATCGCTATTACAGCTATCAA 3542
3480 TGACTATCAGTTCACTTGTAGTCTATCCAGGAAATGATCGCTATTACAGCTATCAA 3539
3543 TTTTGTGGCTACTCTCTGACGAAACAAACAGGATTTGGACATGTTCAATGCCTCCAA 3602
3540 TTTTGTGGCTACTCTCTGACGAAACAAACAGGATTTGGACATGTTCAATGCCTCCAA 3599
3603 GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGA 3662
3600 GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGA 3659
3663 AGAGATGCTGTGTTTCAAAAACCAACATTAAGAGTACAAGATAGTTTCTCTAATGA 3722
3660 AGAGATGCTGTGTTTCAAAAACCAACATTAAGAGTACAAGATAGTTTCTCTAATGA 3719
3723 GAAGTTTGAATTTTGCACACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTTCA 3782
3720 GAAGTTTGAATTTTGCACACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTTCA 3779
3783 CTGGCCCATCAAAATTCAGAT 3803
3780 CTGGCCCATCAAAATTCAGAT 3800

RESULT 3

US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Naele, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21

QY 1787 CCTGTGACCGCTGGTCACTGCTCCAGACCTGATCTCCACATGATGTCAACAGATTG 1846
Db 2053 CTTGTGACCGATGTCAGTGTCTCCAGACCTGAGTCCATCATGATGTCAACAGATTG 2112
QY 1847 GCCATTGACAGTCTTTACACAGACCACTGATGTGTTGGTGGTTTCAATAGTCTCC 1906
Db 2113 GCCATTGACAGTCTTTGACACAGACCACTGATGTGTTGGGCGGTTTCAACAGCTCC 2172
QY 1907 TCCTCAGCGCATCCTGTGATTTACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCG 1966
Db 2173 TCCTCAGTGAAGTCTTGTCTTTACCTCGGAGCAGTGCATGCAACCGCAGTGAAGCTG 2232
QY 1967 CTTGTTTAGCAGAGACCTGATTTGCTGTGTGTGTGGAACACAGGTCGTCTCAGTGA 2026
Db 2233 CTTGTGTGACAGAGACCTGATTCGGTGTCTGTGTGGACACAGTCGTCTCAGTGA 2292
QY 2027 TCTGTGGGCTGCGCACTGATCAACAAAGAAAGTTAAATCAGATGTTTTTCCA 2086
Db 2293 CTTCTGGAGTTGGCACTGAGAAACAGCAGAAAGTTAAATCAGAGTGTTTTCTA 2352
QY 2087 AAAGAACTCTTGAACATGACAGATGTGACAGCAACAGATTTTACAGTGTACAGCCA 2146
Db 2353 AAAGAACTCTTGAACATGACAGATGTGACAGCAACAGATTTTACAGTGTACAGCCA 2412
QY 2147 ACACCAATGACTGCCACTGGTGAATGACCACTGTGTCGCCAGGAACACAGCTGCTCAG 2206
Db 2413 ATACCAATGACTGCCACTGGTGAATGATCACTGTGTCCTGTGAAACACAGCTGCAAG 2472
QY 2207 AAGGCCAGATCTCCATTTTATGATGAGAAATTCGCCAAGGATAACCCCATGCTACT 2266
Db 2473 AAGGCCAGATCTCCATTTGCAAGTATGAGATTTGCCCAAGGATAACCCCATGCTACT 2532
QY 2267 GTAACAGAGACCTGACAGTGTGTCCTGTGACAGCACTGCCAGTGGAGGCC 2326
Db 2533 GCAATAAGAAACCAAGCTGACAGAGTGTGTCCTGTAGACAGAACTGCCAGTGGAGGCC 2592
QY 2327 GGAATCAGAGTGCATTCCTGTCGCCAAATATCTGTGGCAATTTGGTGGCATTTGTTG 2386
Db 2593 GGAATCAGAGTGCATTCCTGTCGCCAAATATCTGTGGCAATTTGGTGGCATTTGTTG 2652
QY 2387 GAAATCTATGTTGAAATTTACTATGCCAAGGAGAAATTAAGCAATGCTAAATGTTCT 2446
Db 2653 GAAATCTGTTGAAATTTACTATGCCAAGGAGAAATTTATGCAATGCTAAATGTTCT 2712
QY 2447 GTAGAAACCAATGCTGCTTTGGTCTCTTACAAACCAAGAGAGTGAATTTGTGCC 2506
Db 2713 GTAGAAACCAATGCTGCTTTGGTCTCTTACAAACCAAGAGAGTGAATTTGTGCC 2772
QY 2507 TTAAGCAGTGGAAATATGCAATCTCAGAGCATCTCAAGCTCACCTTAACCCCAT 2566
Db 2773 TTAAGCAGTGGAAATATGCAATCTCAGAGCATCTCAAGCTCACCTTAACCCCAT 2832
QY 2567 GGGTGGCTTCGGAAGATCAATGTGCTACTGTGCTGGAGAGATGTCCTCCATTTA 2626
Db 2833 GGGTGGCTTCGGAAGATCAATGTGCTACTGTGCTGGAGAGATGTCCTCCATTTA 2892
QY 2627 CAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGTGAATGCTGTGGAATTT 2686
Db 2893 CAATAGTTTGTGAGTGGATGCGCATCTGAGCCAGTGAATGCTGTGGAATTT 2952
QY 2687 TATCAGAACCCAGTACTCGGGAATGAAAGCTGCAACCTGCAATCAACCCATCTCAATGTA 2746
Db 2953 TGTCAAGCTAGTACTCGGGAATTAAGGCTGCAACCTGCAATCAACCCATCTCAATGTA 3012
QY 2747 GTGTGTGAAGGCTGCAACCCAGTGAATGCTGAGCCAGTGGAGATGTCCTGCTTGA 2806
Db 3013 GGGTGTGAAGGCTGCAACCCAGTGAATGCTGAGCCAGTGGAGATGTCCTGCTTGA 3072
QY 2807 GGACAGCATGTGAGATTTACACAGCGGAGCTCTGAGTGCATGTGTTGTCAGCAACATGA 2866
Db 3073 GGACAGCATGTGAGATTTACACAGCGGAGCTCTGAGTGCATGTGTTGTCAGCAACATGA 3132
QY 2867 AGCAGTGTGCACTCCAAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTAGGAAT 2926

Db 3133 AGCAGTGTGCACTCCAAATGCTTACGTGGCTCTTCCCTTTTGGCCAGTGTATGGAAT 3192
QY 2927 GGTATACGATGACACCTGCCCCCTGAAATTTGTTAGGCTACTGTACTCTAGTCAAT 2986
Db 3193 GGTATACGATGACACCTGCCCCCTGAAATTTGTTAGGCTACTGTACTCTAGTCAAT 3252
QY 2987 GCTTGAGCAACACAGCTGTGGTGTGCTGTGATCTCCAGCAATACTGCAAGGGAAT 3046
Db 3253 GCTTGAGCAGCAGCTGTGGTGTGCTGTGATCTCCAGCAATACTGGAAGGAAT 3312
QY 3047 GCATAGAGGTTCTTATAAGGACAGTGAAGATGCTTCCGAAGCCCTACAGGAAT 3106
Db 3313 GTATTGAGGAGCTTATAAGGACCTGTGAAAGTCCGCTCACAGGCTCTGAGGAATG 3372
QY 3107 TGTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGCAGCAGATACACTGT 3166
Db 3373 TGTATCCACAGCCCTTCTGAATCTCCAGCATGTGTCTAGAGCAGCAGATACACTGT 3432
QY 3167 CTTTCAATTCACCTGTCCAGCTTCCAAATGCAACGGCCACAGTAAATGCAATCAAGACA 3226
Db 3433 CTTTCAATTCACCTGTCCAGCTTCCAAATGCAACGGCCACAGTAAATGCAATCAAGACA 3492
QY 3227 TCTGTGAGAGTGTGAGAACTTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCT 3286
Db 3493 TCTGTGAGAGTGTGAGAACTTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCT 3552
QY 3287 TCTACCGTGTATCCCAATGAGGGAATGTCAGCATGCAAGTGAATGGCAGCGCT 3346
Db 3553 TCTATGGTGAACCGCATTAATGAGGCAATGTGAGCATGCAAGTGAATGGCAGCGCT 3612
QY 3347 CTCTGTGCAACACCAACAGCGGCAAGTGTCTTGTGCAACCAAGGCGCTCAAGGGGAG 3406
Db 3613 CACTGTGCAACACCAACAGCGGCAAGTGTCTTGTGCAACCAAGGCGCTCAAGGGGAG 3672
QY 3407 AGTGCAGCTATGTAGGTGAGAAATTCGATCCAGGAAACCTCTCAGAGNACATGTT 3466
Db 3673 AGTGCAGCTATGTAGGTGAGAAATTCGATCCAGGAAACCTCTCAGAGNACATGTT 3732
QY 3467 ATTATCTCTTCTATTGACTATCACTTCACTTCACTTCACTTCACTTCACTTCACT 3526
Db 3733 ACTATACCTTCTCATGTACTATCACTTCACTTCACTTCACTTCACTTCACTTCACT 3792
QY 3527 ATTACACAGCTATCAATTTTGGTGTCTCTGAGCAACAAACAGGATTTGGACATGT 3586
Db 3793 ACTACACAGCTATCAATTTTGGTGTCTCTGAGCAACAAACAGGATTTGGACATGT 3852
QY 3587 TCATCAATGCTCCAGATTTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTG 3646
Db 3853 TCATCAATGCTCCAGATTTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTG 3912
QY 3647 GAAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGATACAAAG 3706
Db 3913 GAACCCAGCTGGAGAGAGTGCCTGTTTCAAAAACCAACATTAAGGATACAAAG 3972
QY 3707 ATAGTTTCTCATGTAGAGTTTGTGTAACCAACCAACCAATATACATTTCTTTGTTT 3766
Db 3973 ATAGTTTCTCATGTAGAGTTTGTGTAACCAACCAACCAATATACATTTCTTTGTTT 4032
QY 3767 ATGTCAGTAAATTTCACTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAAT 3826
Db 4033 ATGTCAGTAAATTTCACTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAAT 4092
QY 3827 TTAAGCAGCTGGTACAGTTCTTCTGAGCTTTCTTCTGAGTGTGTTTCTCTCTCTCTGG 3886
Db 4093 TCATGGACCTGGTACAGTTCTTCTGAGCTTTCTTCTGAGTGTGTTTCTCTCTCTCTGG 4152
QY 3887 TGCTGTGTGTTTGGAGATCAAAACAAAGTTTGGGCTCCAGACGTAGAGAGCAAC 3946
Db 4153 TGCTGTGTGTTTGGAGATCAAAACAAAGTTTGGGCTCCAGACGTAGAGAGCAAC 4212
QY 3947 TTCTTTCAGAGATGCAACAGATGCGCCAGCGCTCCCTTTGCTTGTAAATGTCGCTTGG 4006

```
Db 4213 TTCTTCGGGAGATGCAACAGATGGCCAGCGCCCTTTGCTTCTGTAAACGTTGCTTGG 4272
Qy 4007 AAACAGATGAGAGGCTCTCTGATCTTATTTGGGGGAGATATAAGACTGTTCCCAACCCA 4066
Db 4273 AAACAGATGAGAGGCTCTCTGATCTTATTTGGGGGAGATATAAGACTGTTCCCAACCCA 4332
Qy 4067 TTGCACTGAGGCGGCTTTTGGCAACAAGCGCTGCTCTCTGTTGTTGTGAGGCTCC 4126
Db 4333 TTGCACTGAGGCGGCTTTTGGCAACAAGCGCTGCTCTCTGTTGTTGTGAGGCTCC 4392
Qy 4127 CTCGAGGCTGGGTGGCATCCCTCTCTCTGGGCACTCAGGTCTTGTCTGTGGCCAGCGGCC 4186
Db 4393 CTCGAGGCTGGGTGGCATCCCTCTCTCTGGGCACTCAGGTCTTGTCTGTGGCCAGCGGCC 4452
Qy 4187 TGGTGGACATTTCTCAGCAGATCCCGATAGTGTACAGGAGAGTCAAGGACCGGTGAA 4246
Db 4453 TGGTGGACATTTCTCAGCAGATCCCGATAGTGTACAGGAGAGTCAAGGACCGGTGAA 4512
Qy 4247 ACCGGAAGCAGCAGCGCCCTCTGACAGCCTGGGACCTGCATCTGA 4290
Db 4513 ACCGGAAGCAGCAGCGCCCTCTGACAGCCTGGGACCTGCATCTGA 4556

RESULT 4
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US2002015093A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Query Match 52.4%; Score 2248.2; DB 9; Length 2625;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

Qy 3 GGTGGCCGACGCGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGCGCGGCGGCGGCG 62
Db 9 GGTGGCCGACGCGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGCGCGGCGGCGGCG 68
Qy 63 GCTCGCGGCAAGGAGCGCGGCGCGCACTGGGCACTGGGCACTGGGCACTGGGCACTGGG 122
Db 69 GCTCGCGGCAAGGAGCGCGGCGCGCACTGGGCACTGGGCACTGGGCACTGGGCACTGGG 128
Qy 123 GGGGCTGGGGCGCGGCTGCGCTCCCGCGGCTGCTCTCTCCCGCTGGCGGCGCGCGGCT 182
Db 129 GGGGCTGGGGCGCGGCTGCGCTCTCCCGCGGCTGCTCTCTCCCGCTGGCGGCGCGGCT 188
Qy 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
Db 189 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239
Qy 243 CGAGGCGCGGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Db 240 CGAGGCGCGGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
```

```
Qy 303 TGACCGGCCCTGTGTCAACGGCGGTGCTGTGCAACCTCTGCAACCGGCCAGTGGCTGTGCC 362
Db 300 TGACCGGCCCTGTGTCAACGGCGGTGCTGTGCAACCTCTGCAACCGGCCAGTGGCTGTGCC 359
Qy 363 CGCGGCTGGTGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAACTGGATC 422
Db 360 CGCGGCTGGTGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAACTGGATC 419
Qy 423 TTCTGGTGTGACAGATGGACCTGGAAATATATAAATACAAAAAGAGTGCACGTGGCT 482
Db 420 TTCTGGTGTGACAGATGGACCTGGAAATATATAAATACAAAAAGAGTGCACGTGGCT 479
Qy 483 CATTTGAAGACGACGCAAAATAGATATAGACTTCGTTTTCAATCATTTTGTCTACAGATG 542
Db 480 CATTTGAAGACGACGCAAAATAGATATAGACTTCGTTTTCAATCATTTTGTCTACAGATG 539
Qy 543 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 602
Db 540 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 599
Qy 603 ATTTAGTGGCCTCATTTGTTCTGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTTGC 662
Db 600 ATTTAGTGGCCTCATTTGTTCTGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTTGC 659
Qy 663 CACATCAGGTTATGCTTCTGCTGCAATTTTATGATGCTGCTTATATAATTTGACTGGATT 722
Db 660 CACATCAGGTTATGCTTCTGCTGCAATTTTATGATGCTGCTTATATAATTTGACTGGATT 719
Qy 723 TAATATTACTTACAGTTTGTATGTTGCTCCAAATAACTGCTCAGGCCGAGGAGTGA 782
Db 720 TAATATTACTTACAGTTTGTATGTTGCTCCAAATAACTGCTCAGGCCGAGGAGTGA 779
Qy 783 GATCAGTAATAGCAGCGAACTGTTGAATGTCAATGTTCTGAAAACCTGGAAGGTGAAGC 842
Db 780 GATCAGTAATAGCAGCGGATCTGTTGAATGTCAATGTTCTGAAAACCTGGAAGGTGAAGC 839
Qy 843 ATGTGACATTTCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCACTTGCATTC 902
Db 840 ATGTGACATTTCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCACTTGCATTC 899
Qy 903 AAGTGATGTCAGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCTGATGTTGATTC 962
Db 900 AAGTGATGTCAGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCTGATGTTGATTC 959
Qy 963 TGTACCGACTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAG 1022
Db 960 TGTACCGACTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAG 1019
Qy 1023 AGCATCTCATAAAGCTGTGCTCAATGGAAACATTTATGTGGGTTGTTGAGGATATATGTT 1082
Db 1020 AGCATCTCATAAAGCTGTGCTCAATGGAAACATTTATGTGGGTTGTTGAGGATATATGTT 1079
Qy 1083 CAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCC 1142
Db 1080 CAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCC 1139
Qy 1143 ACTAAACCGTCTCTGTAACAATGTGGTTGTGTAGATATGCTCATTTCTTTGGCAATATACA 1202
Db 1140 ACTAAACCGTCTCTGTAACAATGTGGTTGTGTAGATATGCTCATTTCTTTGGCAATATACA 1199
Qy 1203 GGATAAAATTTACATGTATGGAGGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1262
Db 1200 GGATAAAATTTACATGTATGGAGGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1259
Qy 1263 GAGAGTTTTTCACTTCAATTAATGATCATGGGTGTGTGACCCCTTAAGGCAAAAGGAGCA 1322
Db 1260 GAGAGTTTTTCACTTCAATTAATGATCATGGGTGTGTGACCCCTTAAGGCAAAAGGAGCA 1319
Qy 1323 GTATCGAGTGTGGGCACTCTGCACATTTGTACACTGAGGAATGCGGAGTGGTCAT 1382
Db 1320 GTATCGAGTGTGGGCACTCTGCACATTTGTACACTGAGGAATGCGGAGTGGTCAT 1379
Qy 1383 GCTGCTCATCTTTGGTCACTGCCCTCTCTATGAGATATATAGCAATGTGCAGGAATATGA 1442
```


QY	2954	AAAAATGTT	CAGGCTAC	TGTTAC	TGTTAGT	CAATGTT	TGGAGCA	CAACGAG	CTGTGGCTGGT	3013	
Db	721	AAAAATGTT	TCAGGCTAC	TGTAC	CTGTAGT	CAATGTT	TGGAGCA	CAACGAG	CTGTGGCTGGT	780	
QY	3014	GTACTGAT	CCCCAGCA	TAATCT	CGCAAA	GGGAAATG	CATAGAG	GGTTCCT	TATAAAGAC	CAAG 3073	
Db	781	GTACTGAT	CCCCAGCA	TAATCT	CGCAAA	GGGAAATG	CATAGAG	GGTTCCT	TATAAAGAC	CAAG 840	
QY	3074	TGAAGAT	GCCCTT	CGCAAG	CCCTT	ACAGGAA	ATTTCT	ATCCAC	AGCCCTGCT	CAATTTCCA 3133	
Db	841	TGAAGAT	GCCCTT	CGCAAG	CCCTT	ACAGGAA	ATTTCT	ATCCAC	AGCCCTGCT	CAATTTCCA 900	
QY	3134	GCATGTGT	CTAGAG	GCAG	CAGATAC	AACTGG	CTTTCT	CAATTC	ACTGCTCC	AGCTTGGCCAAT 3193	
Db	901	GCATGTGT	CTAGAG	GCAG	CAGATAC	AACTGG	CTTTCT	CAATTC	ACTGCTCC	AGCTTGGCCAAT 960	
QY	3194	GCAACGG	CCAC	CAGTAA	TGCA	TCAATC	CAGAG	CACTCT	GTGAAAG	TGTGGAACCTCGACCA 3253	
Db	961	GCAACGG	CCAC	CAGTAA	TGCA	TCAATC	CAGAG	CACTCT	GTGAAAG	TGTGGAACCTCGACCA 1020	
QY	3254	CAGCAAG	CACTCG	GAGAC	CTGCA	TATCTGG	CTTCT	ACGGT	GTATCC	CAACATGGAGGGA 3313	
Db	1021	CAGCAAG	CACTCG	GAGAC	CTGCA	TATCTGG	CTTCT	ACGGT	GTATCC	CAACATGGAGGGA 1080	
QY	3314	AATGT	CAGCCA	TGCAAG	TGCAAT	GGGCAC	CGCTCT	CTGTG	CAACACCA	CAACGGGCAAGT 3373	
Db	1081	AATGT	CAGCCA	TGCAAG	TGCAAT	GGGCAC	CGCTCT	CTGTG	CAACACCA	CAACGGGCAAGT 1140	
QY	3374	GCCTT	CTGCA	CCACCA	AGGGCG	TCAAG	GGGGAC	AGTCCC	AGCTAT	GTAGAGTGA	AAATC 3433
Db	1141	GCCTT	CTGCA	CCACCA	AGGGCG	TCAAG	GGGGAC	AGTCCC	AGCTAT	GTAGAGTGA	AAATC 1200
QY	3434	GATACCA	AGGAA	CCCTCT	CAGAG	GAACAT	GTATTAT	TACTCT	CTTATT	TATGACTAT	CAGT 3493
Db	1201	GATACCA	AGGAA	CCCTCT	CAGAG	GAACAT	GTATTAT	TACTCT	CTTATT	TATGACTAT	CAGT 1260
QY	3494	TCACCTT	TAGTCT	ATCCAG	GAAGAT	GCATCG	CTATT	ACAC	AGCTAT	CAATTTT	TGTGGCTA 3553
Db	1261	TCACCTT	TAGTCT	ATCCAG	GAAGAT	GCATCG	CTATT	ACAC	AGCTAT	CAATTTT	TGTGGCTA 1320
QY	3554	CTCTG	CAGCA	CAAAAC	CAGGAT	TTGAC	ATGTTC	ATCA	TGCTCC	AGAAATTT	TCACC 3613
Db	1321	CTCTG	CAGCA	CAAAAC	CAGGAT	TTGAC	ATGTTC	ATCA	TGCTCC	AGAAATTT	TCACC 1380
QY	3614	TCACAT	CACTCG	GGCTGCC	AGTTTCT	CAGCT	GGAA	CCCA	GGCTGG	AGAGAGAT	GCCTG 3673
Db	1381	TCACAT	CACTCG	GGCTGCC	AGTTTCT	CAGCT	GGAA	CCCA	GGCTGG	AGAGAGAT	GCCTG 1440
QY	3674	TTGTTTT	CAAAAA	CCACAT	TAA	GAGTAC	AAAGAT	TATTTCT	CTAT	TATGAGAA	TTTGATT 3733
Db	1441	TTGTTTT	CAAAAA	CCACAT	TAA	GAGTAC	AAAGAT	TATTTCT	CTAT	TATGAGAA	TTTGATT 1500
QY	3734	TTGCG	CAAC	CAAC	CAATAT	CAC	TTTCT	TTGTTAT	GT	CAGTAA	TTTCACTCGCCCATCA 3793
Db	1501	TTGCG	CAAC	CAAC	CAATAT	CAC	TTTCT	TTGTTAT	GT	CAGTAA	TTTCACTCGCCCATCA 1560
QY	3794	AAATTC	AGAT	TGCTTCT	CTCAG	CAC	AGCAATTTT	TATG	GAAC	CTGGTAC	AGTTCTCGTGA 3853
Db	1561	AAATTC	AGAT	TGCTTCT	CTCAG	CAC	AGCAATTTT	TATG	GAAC	CTGGTAC	AGTTCTCGTGA 1620
QY	3854	CTTTCT	TCAG	TGTTTCT	CTTCT	TGCTCG	TGGCTG	CTGTG	TTTTCG	AGATCA	AAAC 3913
Db	1621	CTTTCT	TCAG	TGTTTCT	CTTCT	TGCTCG	TGGCTG	CTGTG	TTTTCG	AGATCA	AAAC 1680
QY	3914	AAAGTT	TGTTGG	GCTCC	AG	CGTAG	AGAC	CACTTCT	TCG	AGAGAT	SCAACAGATGGCCA 3973
Db	1681	AAAGTT	TGTTGG	GCTCC	AG	CGTAG	AGAC	CACTTCT	TCG	AGAGAT	SCAACAGATGGCCA 1740
QY	3974	GCGCT	CCCTT	CGCTCT	GTAA	TGTCG	CTTGAA	ACAGAT	TGAG	AGACCTCT	CGATCTTTA 4033
Db	1741	GCGCT	CCCTT	CGCTCT	GTAA	TGTCG	CTTGAA	ACAGAT	TGAG	AGACCTCT	CGATCTTTA 1800
QY	4034	TTGGG	GGGAG	TATA	AGACT	GTGTTCC	CAAA	CCCAAT	TC	CACTG	GAGCGCTGTTTTGGCAACA 4093

RESULT 6

US-09-893-238-8

: Sequence 8. Application US/09893238

; sequence 8, APPLICATION US
: Patent No. US20020150973A1

; FACEID NO: US20020150
: GENERAL INFORMATION:

APPLICANT: MOORE, K

APPLICANT: MOORE, K.

APPLICANT: NAGIE, D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

1. TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

FILE REFERENCE: 7952-235

FILE REFERENCE: 7853-237
CURRENT APPLICATION NUMBER: 118/08/883 238

; CURRENT APPLICATION NUMBER: US/0
 ; CURRENT FILING DATE: 2001-06-27

; CURRENT FILING DATE: 2001-06-27
 ; PRIOR APPLICATION NUMBER: 08/245,043

PRIOR APPLICATION NUMBER: 09/

;; PRIOR FILING DATE: 1999-02-05
;; PRIOR APPLICATION NUMBER: 50/000 000

;; PRIOR APPLICATION NUMBER: 60/

PRIOR FILING DATE: 1998-07-21

PRIOR APPLICATION NUMBER: 607

; PRIOR FILING DATE: 1998-1

; NUMBER OF SEQ ID NOS: 129

```
; SOFTWARE: Fas
```

; SEQ ID NO 8

; LENGTH: 24

TYPE: DNA

; ORGANISM: Mu

Query Match 35.5%: Score 15218: DB 9: Length 2419.

Query Match 33.3%; SCORE 1321.8
Best Local Similarity 80.7%; Pred No. 0:

Best Local Similarity: 80.7%; Fied: NO: 0;
Matches 1895: Conservative 0: Mismatches 247: Tndels 205: Gaps

09
.
3

QY 3 GG TGG CCG CAG CGG CGG CAA CTG AGG CA AG CTG AGG AGG AGG CCG GCG AG CCG CAG C 62

[illegible]

DD 8Z GG TGG CGC GGC GGT GAG CTT GAG GCG GGT GAG GAC CAG GAG CAG CAG CAG C

I4T

[illegible]

63 GCTCGGGCAGGAGCGGCGGCCGCAC^TGGAAC^TGGAAC^TGACCAAGGCTGGAGGCC 122

143

142 GCCTGGGGCAGGAAGGGCAGGCAGCACCGACCCCTGCACCGCAGGGGCTGGAGGCC 201

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

123 GGGCTGGGGCCGGGCTGGGCTCCCGGGCTGCTGTCTCCACCGCTGGGGCCACGGCT 182

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DB 202 GGGACCGCGGCCCGGCTGTGTCTCCCGGGGTGCTGTCTCGGGCGCTGCCCGCGGCC 261

QY 363 CGCGGCTGGTGGCGAGCAATGCCAGCACTCGGGGGCCGTTTCAG----- 410
Db 433 CACGGCTGGTGGCGAGCAATGCCAGCACTCGGGGGCCGTTTCAGGACATCTGTCTC 492
QY 411 ----- 410
Db 493 ACGCCTAATACACAGCTGTTCCGAAGGTGAGGCTGGAGAAACAGTTTCGAGGCAAGCTTC 552
QY 411 ----- 410
Db 553 GGCTACAGATAAGTTCAAGAGTAACCTGGGGCAACTGGGCTGTCTCCAAAACCAAAA 612
QY 411 ----- 410
Db 613 TGACGAAAAGGACAAAGCTAGACTCTTTTGGGAAAATTTAGCTGACTAATTTTTCACC 672
QY 411 ----ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAAATTATAAATACAAAA 466
Db 673 GAGAACTAACTGGGCTCTCTGGATTTGTACAGATGGACCTGGGAAATTATAATATAAGA 732
QY 467 CGAAGTGCACGTGCTCAATGAAGGACAGGCAAAATAGAAATAGACCTTCGTTTCAATC 526
Db 733 CGAAGTGCACATGCTCAATGAAGGACAGGCAAAATAGAAATAGACCTTCGTTTCAACC 792
QY 527 ATTTTGTACAGATGAGTTGGGACCAATTTATATTTTATGATGGGAGCTCAATTTATG 586
Db 793 ATTTTGTACAGATGAGTTGGGACCAATTTATATTTTATGATGGGAGCTCAATTTATG 852
QY 587 CACCGCTAGTTTGTGCTCAATTTAGTGGCCTCAATTTTCTGAGAGAGATGGCAATGAGACTG 646
Db 853 CACCTCTGATTTGCTGCTTTAGTGGCCTCAATTTTCTGAAAGAGATGGCAATGAGACTG 912
QY 647 TCCGTGAGTTGTGCCACATCAGGTTATGCTTGTCTGCAATTTTATGATGCTGCTT 706
Db 913 CTCCTGAGTCACTGTCACTTCACTTATGATGATGATGATGATGATGATGATGATGATG 972
QY 707 ATAAATTTCACTGGATTTAATATTAATTTAGTGGCCTCAATTTTCTGAGAGAGATGGCAATGAGACTG 766
Db 973 ATAAATTTCACTGGATTTAATATTAATTTAGTGGCCTCAATTTTCTGAGAGAGATGGCAATGAGACTG 1032
QY 767 GCGAGGAGAGTGAAGTACAGTAAATAGCAGGAAACCTGTTGAATGGAATGTTCTGAAA 826
Db 1033 GCGAGGAGAGTGAAGTACAGTAAATAGCAGGAAACCTGTTGAATGGAATGTTCTGAAA 1092
QY 827 ACTGGAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 886
Db 1093 ACTGGAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1152
QY 887 GAGGATCTGCAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 946
Db 1153 GAGGATCTGCAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1212
QY 947 CTGGATGTTCAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1006
Db 1213 CTGGATGTTCAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1272
QY 1007 ACTTAAAGTCTCCAGAGATCTCAATAAAGTGTGGTCAATGGAACCAATTTATGTTGGTGTG 1066
Db 1273 ATTTAAAGTCTCCAGAGATCTCAATAAAGTGTGGTCAATGGAACCAATTTATGTTGGTGTG 1332
QY 1067 TTGGAGGATATATGTTCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1126
Db 1333 TTGGAGGATATATGTTCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1392
QY 1127 CTAGGAGTGGCTTCCACTAAACCGTCTGTGCAACCAATGTTGGTGTGTTAGATATGTTCAAT 1186
Db 1393 CTAGGAGTGGCTTCCACTAAACCGTCTGTGCAACCAATGTTGGTGTGTTAGATATGTTCAAT 1452
QY 1187 CTTTGGCAATTTACAGGATATAATTTACATGATGAGGAGGAAAAATGATCCAACTGGGA 1246
Db 1453 CTTTGGCAATTTACAGGATATAATTTACATGATGAGGAGGAAAAATGATCCAACTGGGA 1512

QY 1247 ATGTGACCAATGAGTTGAGAGTTTTTACATTTCAATAGTCAATGCTGCTGTTGTTGACCC 1306
Db 1513 ACGTGAACCAATGAGTTGAGAGTTTTTCAATTTCAATAGTCAATGCTGCTGTTGTTGACCC 1572
QY 1307 CTAAGGCAAGAGAGAGTATGAGTGGTGGGCACTCTGCAACATTTGTTACCTGAAGA 1366
Db 1573 CGAAAGCTAAGATCAGTATGAGTGGTGGGCACTCTGCAACATTTGTTACCTGGCAT 1632
QY 1367 ATGGCCGAGTGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426
Db 1633 CTGGCCGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1692
QY 1427 ATGTGAGGAAATATGATTTGGAATAAGAACACATGAGTATATTAACACCCAGGTTGCC 1486
Db 1693 TTGTGCGAAATATGATTTGGAATAAGAACACATGAGTATATTAACATACTAGGTTGCTC 1752
QY 1487 TTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTGACCATAGGACCGGCTTATAGC 1546
Db 1753 TTGTGCAAGGGGTTATGGCCACAGTAGTGTGTTGATGACAGGACCAAGGCTCTGTAGC 1812
QY 1547 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606
Db 1813 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1872
QY 1607 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1666
Db 1873 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1932
QY 1667 TGCAACACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1726
Db 1933 TGCAACACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1992
QY 1727 ACACATCTATGAGCCATGCGCCAAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1786
Db 1993 ACATTTCCATGAGCCACCGGTGCGCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2052
QY 1787 CTTGTGACCGTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACACAGATTTG 1846
Db 2053 CTTGTGACCGTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACACAGATTTG 2112
QY 1847 GCAATTCAGCAGCTTTACACACAGCACCATGATGATGATGATGATGATGATGATGATGATG 1906
Db 2113 GCAATTCAGCAGCTTTACACACAGCACCATGATGATGATGATGATGATGATGATGATGATG 2172
QY 1907 TCTCAGCGACATCTCTGTTATTCACCTCGAAACAGTGTGATGCGCATCGAGTGAAGCG 1966
Db 2173 TCTCAGCGACATCTCTGTTATTCACCTCGAAACAGTGTGATGCGCATCGAGTGAAGCG 2232
QY 1967 CTTGTTTAGCAGCAGGACCTGTTATTCGGTGTGTTGGAAACACAGGCTGCTCTCAGTGTA 2026
Db 2233 CTTGTTTAGCAGCAGGACCTGTTATTCGGTGTGTTGGAAACACAGGCTGCTCTCAGTGTA 2292
QY 2027 TCTCGTGGCGCTGGCAACTGATGAACAAAGAAAAAGTTAAATCAGAAATGTTTTTCCA 2086
Db 2293 CTTCTGGAGTTGGCAACTGATGAACAAAGAAAAAGTTAAATCAGAAATGTTTTTCCA 2352
QY 2087 AAAGAACTTTGACCATGACAGATGTGACCAAGCAGACAGATTTGTTACAGCTGTACAGCCA 2146
Db 2353 AAAGAACTTTGACCATGACAGATGTGACCAAGCAGACAGATTTGTTACAGCTGTACAGCCA 2412
QY 2147 ACACCAA 2153
Db 2413 ATACCAA 2419

RESULT 7
US-10-197-824-1
; Sequence 1, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
; FILE REFERENCE: 9U 106 R1

/ CURRENT APPLICATION NUMBER: US/10/197,824

/ CURRENT FILING DATE: 2002-07-19

/ NUMBER OF SEQ ID NOS: 39

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 1

/ LENGTH: 6733

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: (575)...(4711)

/ OTHER INFORMATION:

US-10-197-824-1

Query Match 34.1%; Score 1463.8; DB 16; Length 6733;

Best Local Similarity 62.9%; Pred. No. 0;

Matches 2490; Conservative 0; Mismatches 1377; Indels 90; Gaps 11;

QY	332	GCAACCCCTGGCCGCGCGAGTCGCTGCCCGCGCGCTGGTGGCGAGCAATGCCAGC	391
DB	789	GCCGCTGTCAACTCCACCTGCCCTCGGACCCGCGCTGGTGGCGAGCAATGCCAGC	848
QY	392	ACTGGGGGGCGCTTCAGACTTAACTGGATCTCTCGGCTTTGTGACAGATGGACTGGAA	451
DB	849	ACTGCCAGGCGAGTTCAAGTTTACAGAACTTCTGGATATTTAACAGATGCCCAATTA	908
QY	452	ATTATAAATACAAACGAGTGCAGTGGCTCATTTGAAGGACAGCAATAGAAATAGA	511
DB	909	ACTATAAATATAAACTAAATGTACTTGGCTCAITGAAGGCTATCCAAATCCAGTGTAA	968
QY	512	GACTTCGTTCAATCAATTTGTACAGAGTGTAGTTGGACCAITTTATATGTTTATGATG	571
DB	969	GATTAAAGATTCAATCAATTTGTCTACAGAACTGTAGTGGATCATATGATGTTTATGATG	1028
QY	572	GGGACTCAATTTATGACCGGTAGTGTGTCATTTAGTGGCTCATTTGTTCCCTGAGAG	631
DB	1029	GAGATTCAATATATGACCTTTAATAGCTGTACTTAGTGGTTTGATAGTCCCTGAAATAA	1088
QY	632	ATGGCAATGAGACTGTCCCTCAGGTTGTGGACATCAGTGTATGCTTGTGCTATTTT	691
DB	1089	GGGGCAATGAAGTGTGCTGAGTTGTACTACATCTGGCTATGCACTGTTCATTTT	1148
QY	692	TTAGTGATGCTGCTTAAATTTGACTGGAATTAATTTACTTACAGTTTGTATGATGTC	751
DB	1149	TTAGTGATGCTGCTGATTAATCTAACTGGTTTCAACATTTTCTATTCAATCAATTTTGTGTC	1208
QY	752	CAATAACTGCTCAGGCGGAGAGGTAGATCAGTAAATAGCAGCAACT---GTTG	808
DB	1209	CTAACAAATGCTGCTGTCATGGAGGTGACAACTAGTGTCTGTGTTCCCAAGTCAAGTAT	1268
QY	809	AATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACA	868
DB	1269	ATTGTGAATGTGAATAATCTGGAAGGTGAAGCTTGTGATATTTCTTACTGTAAAGCCA	1328
QY	869	ACTGTGTTTTCCTCATCGAGCATCTGCAATTCAGTATGTCAGAGGATGCTCCTGCT	928
DB	1329	ATTGGCGAGTCCAGATCACGGTTACTGTGACCTGACTGGAGAAAATTTATGTGTGCA	1388
QY	929	TCTCAGACTGGCAGGCTCCTGGATGTTCAAGTCTCTGTACCAAGCTTAACCACTATTGGA	988
DB	1389	ATGATAGTTGGCAAGGTCCTGATTTCTTTTGAATGTTTCCCTTACTGAGTCTTACTGGA	1448
QY	989	CTCAGAGGAATATCTTAACCTAAAGC-----TCCCCAGAGCATCTCAATAAGCTGTGG	1042
DB	1449	TTCTGCCAAAGCTTAAACCTTCACTGCTCTGTAGGTGCGGCTTACATAAAGCAGTTT	1508
QY	1043	TCAATGGAACATTTATGTGGTGTCTGAGGATATGTTCAACCACTCAGATATAACA	1102
DB	1509	TACAGGGAATTTATGTGGTGTGATTTGGATATCTTTTAACTACAGTCTTTTCAAA	1568
QY	1103	TGGTTCTAGCGTATGACTTGTCTTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACA	1162
DB	1569	TGGTCTTAAATTAACAATTTAGAAAGCAGTATATGGAATGTAGGAATCTCCATCAAGGGAC	1628

QY	1163	ATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTTACATGATG	1222
DB	1629	CT---CTCCAGAGATATGGACACTCTCTTGGCTTTATATCAGGAACAATCTTTATGATG	1685
QY	1223	GAGGAAAAATTTGATCCAA---CTGGGAATGTGACCAATGAGTTGAGTTTTCACATTC	1279
DB	1686	GAGGCAGAAATTTGAACAAATGATGGCAATGTACAGATGAATTTATGGTCTTTTAAACATAC	1745
QY	1280	ATAATGAGTCATGGGTGTTGTTGACCCCTAAAGCAAGG-----AGCAGTAGCAG	1330
DB	1746	ATAGTCAGTCATGAGTACAAAACCTCTACTGTTCTTGACATGGTCAGCAGTATGCTG	1805
QY	1331	TGTTTGGGCACTCTGCACACATTTTACACTGAGAATGCCGAGTGGTCATGCTGGTCA	1390
DB	1806	TGGAGGGACATTCAGCACATATTTATGGAGTTGATAGTAGATGTTGTCAATGATCATAA	1865
QY	1391	TCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGAGGAATATGATTTGGATA	1450
DB	1866	TATTTGGATATTTCTGCAATATATGTTTATACAGCAGCATACAGGAATACCATATCTCAT	1925
QY	1451	AGAACACATGGAGTATATTTACACACCCAGGGTGCCCTTGTCAAGGGGGTTACGGCCATA	1510
DB	1926	CAAAACATTTGGCTTGTTCAGAAAATTAAGGAGCTATTGTACAAGTGGATATGGCCATA	1985
QY	1511	GCAGTGTTTACGACCATAGGACCGAGGCCCTATACGTTTATGTTGGTGGCTACAGGCTTTCA	1570
DB	1986	CTAGTGTGTATGATGAATTAACAAAGTCCATTTATGTTTATGAGGTTTAAAGCATTTGC	2045
QY	1571	GTGCCAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGATACCCAGATGT	1630
DB	2046	CAGGGAACAATATGATTTGGTTGTATGATCTTTTAAATATGAAGTTTAACTAGACTT	2105
QY	1631	GGACCACTTTAAGGACCGGATTTTCCGTTACTTGCACACAGCTGTGATGAGTGTG	1690
DB	2106	GGACTATTTTGAAGAAAGTGGTTTGCAGATACCTTCATTTCAGCTGTCTTATCAATG	2165
QY	1691	GAACCATGCTGTTGTTGGGGGAAACACACACAATGACATCTATGAGCCATGGCCCA	1750
DB	2166	GAGCTATGCTTATTTTGGAGAAATACCCATAATGACACTTCTTGGATTAACGGTGC	2225
QY	1751	AATGCTTCTCTTTCAGATTTTATGGGCTATGACATTTGCGCTGTGACCGCTGGTCA	1810
DB	2226	AATGTTTTCTCCGATTTCTTGGCATATGACATAGCTTGTGATGAATGGAATTA	2285
QY	1811	CCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGACGCTTACACACA	1870
DB	2286	CAAAACCAATCTTCATAGATGTCAACAGATTTGGACACTCTGTCAGTAGTCAATTA	2345
QY	1871	GCACCATGATGTTGTTGCGGTGTTTCAATAGTCTCTCTCCTCAGCCACATCTGTTAT	1930
DB	2346	GGTCCATGATATATTTTGGGGGATTTTCTAGTGTACTCTTAAATGATATCTTGTATACA	2405
QY	1931	CCTCGGAACAGTGTGATGGGCATCGGAGTGAAGCGCTTTGTTAGCAGACAGACCTG	1990
DB	2406	AGGCTCCAAATTTGCAAGGCTTTTACAGATGAAGAACTTTGTAAAAAATGCTG	2465
QY	1991	TTCCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCTGGGCGCTGCCA	2050
DB	2466	TAAATGTTGTTGGAAATAAATCACTGTGAAT-----CTTGGGAATCTGGGA	2513
QY	2051	AACAAAGAAAGTTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATCA	2110
DB	2514	ATACTAATAATATTTCTAGAGCAAAAGTCCCTCTCTTAAACAGCTCTTCTGATG	2573
QY	2111	GTGACCAACACACAGATTTTACAGCTGTACAGCCCAACCAATGACTGCGCAC	2170
DB	2574	GTTACAGATATGCAATTTGCGCAGCTGTACTGCCAATACAAATGGGTGGCCAT	2633
QY	2171	ATGACCAATGTTGTCCCCAGGAACCAACAGCTGCTCAGAAAGGCCAGATCTCC	2230
DB	2634	ATGCAAGAAATGCAATTTCCGGCAATAGTAACCTGCAAGTATGTTCTGTCAAGAA	2693

2231 ATGAGAAATGGCCCAAGGATAACCCCATGTACTACTGTACACAGAGAGACCCAGCTGCAGGA 2290
Db
2694 AATGTCATGTGAGAAATGA-----GCAGATTTGTAAACAACTTACAGCTGTAAAA 2744
Qy
2291 GCTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGAATGCCCTGC 2350
Db
2745 GCTGTCACTAAACTTGAATTTGCCAGTGGGATCAGACAGCAGGAATGCCAGGCTTTAC 2804
Qy
2351 CCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAATCATGTTTGAATTAATA 2410
Db
2805 CAGCTCATCTTTTGGAGAGGATGAGATCATATTTGGGGATGCTTGTCTTAGAGTCAAT 2864
Qy
2411 CTGCAAGAGGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAATGCCCTTTTGG 2470
Db
2865 CCAGTAGAGAACTATGACAAATGCTAAATGTTCTGTAGGAACCAATGCCCTTTTGG 2524
Qy
2471 CTCTCTTCAACCCAGAGAGTAGAATTTGCTTAAAGCAGCTCCGGAATTAATGAGT 2530
Db
2925 CTTCAATTAACAACTTCAAAAGAGTAGAATTTGTC-----TGGATG 2966
Qy
2531 CATCTCAGAGCATGTCCAAGTCACTTAAACCCATGGTGGCTTTCGGAAGATCAATG 2590
Db
2967 AAATACAGAGTATACACACAGAAAGTATCACCTTGGTAGGCTTGGCAAGATCAATA 3026
Qy
2591 TGTCTACTGTGTGGGAAGATATGCCCAATTAATAATAGTTTACTACAGTGAATGC 2650
Db
3027 TATCTATTTGGGATGGAGACATGCTCTCTTTTCAACACCACTACAGTGGCTTC 3086
Qy
2651 CGTCTGACCCAGTACTGCTGATCTGTGGAATTTTACAGAACCCAGTACTCGGGAC 2710
Db
3087 CTGGCGAACCCAAATGATCTGGGTTTGTGGCATATCTGGAAGGCTGCGATGGCGCT 3146
Qy
2711 TGAAGGCTGCAACCTGCATCAACCCACTCAATGATGATGCTGTGAAAGGCTG----- 2764
Db
3147 TAAAGCTAATCTTGTACATCTATGCAATGSCCTTGTCTGTGAAACCTTGTGTTA 3206
Qy
2765 ---CAAACCAAGTGTAAAGCAGTCCGGAACACATGTCCTTGAGACAGCATGTGGAG 2821
Db
3207 GTCCAAATCAAAATCGGAGGCGTGCAAAAGCCATGCTCTCTGAGACATCATGTTCCA 3266
Qy
2827 ATTGCACAGGCGAGCTCTGAGTGCATGTGTCAGCAACATGAAGCAGTGTGAGCT 2881
Db
3267 ACTGTACAGCAATGCGATGAGTGTATGTGTGAGCAGCATGACAAAGATGTGTGACT 3326
Qy
2882 CCAATGCTATGTGGCTCTCCCTTTTGGCCAGTGTATGGAATGTATACAGTAGCA 2941
Db
3327 CTAAATGCTATATCATCTCTTTTCCATATGACAAATGTAGAGTGGCAAC---TGCCA 3383
Qy
2942 CTTGCCCTTGAATAATTTGTTGAGCTACTGATCTGTAGTCAATGCTTGGAGCAACAG 3001
Db
3384 CTTGCTCCCTCTAAATTTGTTGAGTGAACCTGTGGAAGTGTGGAAGCTGTTTGAACAG 3443
Qy
3002 GCTGTGGCTGTGTACTGATCCAGCAATATGCGCAAGGGAATGCAATAGAGGTTTCT 3061
Db
3444 GATGTGGCTGTGCAATGATCTAGTATACAGGAAGAGACATTTGCAATGAAGTTCTT 3503
Qy
3062 ATAAAGACAGTGAAGTGTGCGAAGCCCTTACAGGAATTTCTATCCAGGCC 3121
Db
3504 CACGGGACCAATGAAG-----CTTATTTGGAATGCCACCAATGAGATGG 3548
Qy
3122 TGCTCAATTTCCAGCATGTGTGTAGAGGACAGCATACAACTGCTTTTCAATTCATCTG 3181
Db
3549 TTTCTGACACCAATCTTTTGCCCAAGAAAGAACTATGAGTGTCTTTATCCAGTGT 3608
Qy
3182 CAGCTTGCCAATGCAACGCCACAGTAATATGATCAATCAGAGCATCTGTGAGAGTGTG 3241
Db
3609 CAGCTTGCCAGTGAATGACATAGCACTTGCATCAATAATAATGTGTGCGAAGTGTGA 3668
Qy
3242 AGAATCTGACACAGGACAGTCTGAGACCTGCATATCTGGCTTCTAGCGTATCCCA 3301
Db
3669 AAAATCTCACCACAGGAAGAGTGTCAAGATTTGATCCAGGTTATTTAGGATCAAA 3728
Qy
3302 CCAATGGAGGAAATGTGAGCCATGCAAGTGAATGCGACGCTCTCTGTGCAACCA 3361

3729 CCAATGGTGGACAGTGCACAGCTTGTACATGTCAGTGGCCATGCAAAATATCTGTCTAT 3788
Qy
3362 ACACGGGCAAGTGTCTTTCACCAACCAAGGGGTCAAGGGGACGAGTGGCCAGCTATGTG 3421
Db
3789 ACACAGGAAATGTTTTCACCACTAAAGGAATAAAGGTTACCAATGCCAAATATGTG 3848
Qy
3422 AGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACAATGTTATTATATCTTCTTA 3481
Db
3849 ACTCTGAAATCGCTATGTTGTAATCCACTTAGAGGAACAATGTTATTACAGCTTTGA 3908
Qy
3482 TTGATATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGATATCA 3541
Db
3909 TTGATATCAATTTACCTTTAGCTTTATTACAGGAAGATGATCGCCACCATATCTGCATA 3968
Qy
3542 ATTTTGTGCTACTCTCTGACGAACAAACAGGATTTTGACATGTTTCATCAATGCTCCA 3601
Db
3969 ACTTTATAGCAACCCAGACAGTGCACAAATCTGGATATATCAATTAATGATCAA 4028
Qy
3602 AGAATTTCAACCTCAACATCACCTGGGTGTCAGTTTCTCAGCTGGAACCCAGGCTGGAG 3661
Db
4029 ACACTTAACTCAACATTTCTGCTGCTGTCGGTTCAACAGCTGGAACAATATCTGGGG 4088
Qy
3662 AAGATGCTCTGTTTTCANAAACCAATTAAGAGTACAAAGATAGTTTCTCTAATG 3721
Db
4089 AAGAGACTTCTATAGTTTCCAGAAATAATAAAGAAATACAGATAGTTTTCCTATG 4148
Qy
3722 AGAAGTTTCATTTTGGCAACCCCAATATACATTTCTTTTGTATGTACGATTTCA 3781
Db
4149 AAAATTTAACTTTAGAGCAATCTCAATTTATCTGTGTAGCTGACCACTTT 4208
Qy
3782 CTTGGCCCATCAAAATTCAGATGCTCTCTCAGACAGCAATTTTATGACCTGGTAC 3841
Db
4209 CTTGGCTATTAAATACAGATTTGATTTCTCAACACAAATCAATCATGAGCTGTGCG 4268
Qy
3842 AGTTCTCTGCTGCTTCTTCACTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3901
Db
4269 AGTTTCTGCTGCTTCTTCACTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4328
Qy
3902 GGAAGATCAAAACAAAGTGTGGGCTCCAGACGTAGAGAGCAATCTTTCAGAGATGC 3961
Db
4329 GGAAGATCAAAACAACTTGTGGGCTTCTCGACGGAGAGCACTGCTTTCAGAACGAC 4388
Qy
3962 AACAGATGCCAGCGTCT 4021
Db
4389 AGCAGATGCCAGCGTCT 4448
Qy
4022 CTCTCATCTTATTGGGGGAGTATAAAGACTGTTTCCCAACCCATTTGCACTGGAGCCCT 4081
Db
4449 AACAGAGTTCTGCGAGGCGCATTAGAGGGGCGACCCCAAGCCAAATTGCCATTGAACCA 4508
Qy
4082 GTTTTGGCAACAAAGCCGCTGCTCTCTCTGTTTGTAGGCTCCCTCGAGGCTGGGTG 4141
Db
4509 GTCTGGGAACAGAGCTGCTGTTCTGACTGTGTTTCTTTGTACCAAGGATCATCAG 4568
Qy
4142 GCATCCCTCTCTCTGGGAGTCAAGGCTTCTGCTGGCCAGCGCTCTGGTGGACATTTCTC 4201
Db
4569 GTGCCCTCTCCCTGGGAGTCAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4628
Qy
4202 AGCAGATGCCGATGATGTAACAGGAAGTCAAGGAGCCCTGAGAAACCCGAGAGCAGC 4258
Db
4629 AACAGAAAGCTTCAGATAGTAAAGATAAGACTTCTGGAGTCCGGATCGAAACACC 4685

RESULT 8

US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

FILE REFERENCE: 7853-237
CURRENT APPLICATION NUMBER: US/09/893,238
CURRENT FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: 09/245,041
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/093,630
PRIOR FILING DATE: 1998-07-21
PRIOR APPLICATION NUMBER: 60/104,978
PRIOR FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 129
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 1051
TYPE: DNA
ORGANISM: Mus musculus
US-09-893-238-10

Query Match 14.9%; Score 640.2; DB 9; Length 1051;
Best Local Similarity 86.1%; Pred. No. 6.2e-176;
Matches 724; Conservative 0; Mismatches 108; Indels 9; Gaps 1;
QY 3 GGTGCGCGCAGCGCGGCAACTGAGGCAAGGTGAGGAGGAGGAGCGGCGGAGCGAGC 62
DB 216 GGTGCGCGCGCGCGGCGACTGAGCGCGGCTGAGGCGGAGGAGGAGCGAGCAGCAGC 275
QY 63 GCTCGCGGCGAGGAGCGCGCGGCGGCGCACTGGGACTGGGACCTGACACAGGCTGGGAGGCC 122
DB 276 GCGTCGCGGCGAGGAGGCGAGGCGAGCAGCAGCCCTGACCGCGAGGCGGCTGGAGGCC 335
QY 123 GGGGCTGGGGCGCGGCTGCGCTCCCGCGGCTGCTCTCCACCGCTGCGGCGACCGCT 182
DB 336 GGGACCGCGCGCGCGGCGGCTGCTCTCCCGCGGCTGCTGCGGCGGCGCTGCCCGCGGCC 395
QY 183 GCTGCTGCTGCTGTTGTTGCTCCCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGAGGC 242
DB 396 GCTGCTGCGGCTGCTCTT-----TTGCTGCTGCTGCTGCTGCTGCTGCTGCGCGGAGGC 446
QY 243 CGAGCGCGCGCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGGC 302
DB 447 CGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506
QY 303 TGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
DB 507 TGACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566
QY 363 CGCGCGCTGGGTCGCGAGCAATGCGAGCACTGCGGCGCGCGCTTCAAGCACTGAGATC 422
DB 567 CACGCGCTGGGTCGCGAGCAATGCGAGCACTGCGGCGCGCGCTTCAAGCACTGAGATC 626
QY 423 TTCTGGGTTTGTGACAGATGACCTGGAAATTTATAATACAAACGAGTGCAGTGGCT 482
DB 627 TTCTGGGTTTGTGACAGATGACCTGGAAATTTATAATACAAACGAGTGCAGTGGCT 686
QY 483 CATTGAAGGACAGCGCAATAGAAATAGAAATTTGTTCAATCATTTTGTACAGATG 542
DB 687 CATTGAAGGACAGCGCAATAGAAATAGAAATTTGTTCAATCATTTTGTACAGATG 746
QY 543 TAGTTGGGACCATTTATATTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 602
DB 747 TAGTTGGGACCATTTATATTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 806
QY 603 ATTTAGTGGCTCATTTGTTCTGAGAGAGATGGAATGAGACTGCTCCCTGAGGTTGTTGC 662
DB 807 CTTTAGTGGCTCATTTGTTCTGAGAGAGATGGAATGAGACTGCTCCCTGAGGTTGTTGC 866
QY 663 CACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
DB 867 CACTTCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
QY 723 TAATATTACTTACAGTTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
DB 927 TAATATTACTTACAGTTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986

QY 783 GATCAGTAATAGCAGCAAACTGTTCAATGTGAATGTTCTGAAAACTGGAAGGTGAAGC 842
DB 987 GAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAAATGTTCTGAAAACTGGAAGGCGCG 1046
QY 843 A 843
DB 1047 A 1047

RESULT 9

US-10-198-846-5869/c
Sequence 5869, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5869
LENGTH: 625
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
OTHER INFORMATION: n = A, T, C or G
US-10-198-846-5869

Query Match 6.5%; Score 277.4; DB 14; Length 625;
Best Local Similarity 99.6%; Pred. No. 7.1e-70;
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2699 GTACTCGGGGACTGAAGGCTGCAACTGCATCAACCCACTCAATGCTAGTGTCTGAAA 2758
DB 340 GTACTCGGGGACTGAAGGCTGCAACTGCATCAACCCACTCAATGCTAGTGTCTGAAA 281
QY 2759 GGCCTGCAAAACACAGTGTCTAAGCAGTCCCGACACCATGTGCTTGGAGCAGCATGTG 2818
DB 280 GGCCTGCAAAACACAGTGTCTAAGCAGTCCCGACACCATGTGCTTGGAGCAGCATGTG 221
QY 2819 GAGATTGCACAGCGGCGCTGAGTGCATCTGCTGAGCAACATGAAGCAGTGTGTTGG 2878
DB 220 GAGATTGCACAGCGGCGCTGAGTGCATCTGCTGAGCAACATGAAGCAGTGTGTTGG 161
QY 2879 ACTCCAATGCCCTATGTGGCCTCTTCCCTTTTGGCAGTGTGATGGAATGATACGATGA 2938
DB 160 ACTCCAATGCCCTATGTGGCCTCTTCCCTTTTGGCAGTGTGATGGAATGATACGATGA 101
QY 2939 GACCTGCCCCCTGAAAATTTTTCAGGCTACTGTACCT 2977
DB 100 GACCTGCCCCCTGAAAATTTTTCAGGCTACTGTACCT 62

RESULT 10

US-09-864-761-15900
Sequence 15900, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17

; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10

; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 966
; LENGTH: 246

; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-902-941-966

Query Match 5.7%; Score 246; DB 9; Length 246;

Best Local Similarity 100.0%; Pred. No. 5.5e-61;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 2574

DB 246 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 2634

DB 186 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

RESULT 13

US-09-849-626-966/c

; Sequence 966, Application US/09849626

; Publication No. US20020197669A1

; GENERAL INFORMATION:

; APPLICANT: Bangur, Chaitanya

; APPLICANT: Fanger, Gary

; APPLICANT: Wang, Aijun

; APPLICANT: Wang, Tongtong

; APPLICANT: Switzer, Anne

; APPLICANT: McNeill, Patricia

; APPLICANT: Clapper, Jonathan

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C16

; CURRENT APPLICATION NUMBER: US/09/849,626

; CURRENT FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 1926

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-849-626-966

Query Match

Best Local Similarity 100.0%; Pred. No. 5.5e-61;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 2574

DB 246 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 2634

DB 186 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

DB 186 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

RESULT 14

US-10-017-754-966/c

; Sequence 966, Application US/10017754

; Publication No. US20030054363A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Carter, Darick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C18

; CURRENT APPLICATION NUMBER: US/10/017,754

; CURRENT FILING DATE: 2001-10-29

; NUMBER OF SEQ ID NOS: 2004

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-017-754-966

Query Match

Best Local Similarity 100.0%; Pred. No. 5.5e-61;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 2574

DB 246 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 2634

DB 186 CTTGCGAAGATCAATGTGCTCTTACTGTGCTGGGAAGATATGTCCTTAAACCAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

RESULT 15

US-10-113-872-966/c

; Sequence 966, Application US/10113872

; Publication No. US20030170255A1

: GENERAL INFORMATION:

APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.

APPLICANT: Henderson, Robert A.

APPLICANT: Kallos, Michael D.

APPLICANT: Sleath, Paul R.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Fander, Gary R.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C19

FILE REFERENCE: 22-11147001
: CURRENT APPLICATION NUMBER: US/10/113,872

CURRENT AFFILIATION NUMBER: 007-1
: CURRENT FILING DATE: 2002-03-28

CURRENT FILING DATE: 2002
NUMBER OF SEC ID NOS: 2011

```

; NUMBER OF SEQ ID NOS: 2011
: SOFTWARE: FastSeq for Windows Version 4.0

```

; SOFTWARE: TASC
: CEO ID NO 966: SEQ ID NO 566
: LENGTH: 246

```

; LENGTH: 2
TYPE: DNA

```

TYPE: DNA
ORGANISM: Homo sapiens

; ORGANISM: HOMIO
IIS-10-113-873-966

```
Query Match          5.7%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 5.5e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	2515	CTGCGAATAATGCAGTCAATCTCAGAGCATGTCCAAAGCTACCTTAACCCCAATGGGTGGCG	2574
Ab	2445	CTGCGAATAATGCAGTCAATCTCAGAGCATGTCCAAAGCTACCTTAACCCCAATGGGTGGCG	187

246 CTGGGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGGC 187

2575 CTTCCCAAGATCAATCTGTCCTACTGCTGCTGGGAAGATATGTCCCCATTACAAATAGT 2634

106 :GTTGGCGAGTCTGTCTCCGACATATCTCAAAATAGT 127

[illegible]

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

THE

QY 2155 GAAAGG Z

Search completed: March 5, 2004. 13:18:46

```
search completed: March
Job time : 962.325 secs
```

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	4184.8	87.5	8589	3	US-09-245-041-14	Sequence 14, Appl
2	3696.2	96.2	4072	3	US-09-245-041-16	Sequence 16, Appl
3	3398	79.2	8827	3	US-09-245-041-1	Sequence 1, Appli
4	2248.2	52.4	2625	3	US-09-245-041-18	Sequence 18, Appl
5	1853	43.2	6370	3	US-09-245-041-12	Sequence 12, Appl
6	1521.8	35.5	2419	3	US-09-245-041-8	Sequence 8, Appli
7	640.2	14.9	1051	3	US-09-245-041-10	Sequence 10, Appl
C 8	246	5.7	246	4	US-09-702-705-966	Sequence 966, App
C 9	246	5.7	246	4	US-09-736-457-966	Sequence 966, App
C 10	246	5.7	246	4	US-09-614-124B-966	Sequence 966, App
C 11	246	5.7	246	4	US-09-671-325-966	Sequence 966, App
12	212.8	5.0	415	4	US-09-833-381-1916	Sequence 1916, Ap
13	201.2	4.7	9050	3	US-09-245-041-5	Sequence 5, Appli
14	169.2	3.9	17056	3	US-09-245-041-3	Sequence 3, Appli
C 15	132.2	3.1	226	3	US-09-245-041-6	Sequence 6, Appli
C 16	115.4	2.7	221	3	US-09-245-041-7	Sequence 7, Appli
17	100.2	2.3	5973	3	US-09-245-041-4	Sequence 4, Appli
C 18	72.8	1.7	477	3	US-09-135-994-1	Sequence 1, Appli
C 19	72.8	1.7	477	4	US-09-684-843A-1	Sequence 1, Appli
C 20	72.6	1.7	10348	2	US-08-457-273B-41	Sequence 41, Appl
C 21	72.6	1.7	10348	3	US-08-556-41B-13	Sequence 13, Appl
C 22	72.6	1.7	10348	3	US-09-041-886-14	Sequence 14, Appl
C 23	72.6	1.7	10366	1	US-08-246-982A-5	Sequence 5, Appli
C 24	72.6	1.7	10366	1	US-08-453-265-5	Sequence 5, Appli
25	70.4	1.6	6530	2	US-08-146-630-1	Sequence 1, Appli
26	70.4	1.6	6530	3	US-08-458-240-1	Sequence 1, Appli
27	70.4	1.6	6530	5	PCT-US93-03993-1	Sequence 1, Appli


```
QY 2163 CTGCTGCAATGACCATTTGTCTCCAGGAACACAGCTGCTCAGAAAGGCCAGATCTCCAT 2222
Db 2160 CTGCTGCAATGACCATTTGTCTCCAGGAACACAGCTGCTCAGAAAGGCCAGATCTCCAT 2219
QY 2223 TTTTAGTATGAAATTTGCCCAAGGATTAACCCCATGTAATCTGTAAACAGAGACCAG 2282
Db 2220 TTTTAGTATGAAATTTGCCCAAGGATTAACCCCATGTAATCTGTAAACAGAGACCAG 2279
QY 2283 CTGCAAGAGCTGTGCCCTGACAGCAATCTCCAGTGGAGCCCGGAAATCAGAGTGCAT 2342
Db 2280 CTGCAAGAGCTGTGCCCTGACAGCAATCTCCAGTGGAGCCCGGAAATCAGAGTGCAT 2339
QY 2343 TGCCTCCGCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAATCTCATGTTTGA 2402
Db 2340 TGCCTCCGCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAATCTCATGTTTGA 2399
QY 2403 AATTACTACTGCCAGAGAAATTTAGCAATGCTTAATTTCTGTAGGAACACCAATGC 2462
Db 2400 AATTACTACTGCCAGAGAAATTTAGCAATGCTTAATTTCTGTAGGAACACCAATGC 2459
QY 2463 CCTTTGGCTTCTCTTACACCCAGAAAGGTAGAAATTTGTCTTTAAGCAGCTGCGAAT 2522
Db 2460 CCTTTGGCTTCTCTTACACCCAGAAAGGTAGAAATTTGTCTTTAAGCAGCTGCGAAT 2519
QY 2523 AATGAGTATCTCAGAGCATGTCCTAAGCTCACTTAACCCCATGGTCCGCCCTTCGAA 2582
Db 2520 AATGAGTATCTCAGAGCATGTCCTAAGCTCACTTAACCCCATGGTCCGCCCTTCGAA 2579
QY 2583 GATCAATGTGCTTACTGCTGGAGATATCTCCCAATTTTACAAATAGTTTACTACA 2642
Db 2580 GATCAATGTGCTTACTGCTGGAGATATCTCCCAATTTTACAAATAGTTTACTACA 2639
QY 2643 GTGATGCCGCTCAGAGCCAGTGTGCTGGAAATCTGTGGAATTTTATCAGAACCCAGTAC 2702
Db 2640 GTGATGCCGCTCAGAGCCAGTGTGCTGGAAATCTGTGGAATTTTATCAGAACCCAGTAC 2699
QY 2703 TCGGGGACTGAGGCTGCACTGCTCACTCAACCCACTCAATGGTAGTCTGTGAAAGGCC 2762
Db 2700 TCGGGGACTGAGGCTGCACTGCTCACTCAACCCACTCAATGGTAGTCTGTGAAAGGCC 2759
QY 2763 TGCACACACAGTGTCTAAGCAGTCCGACACCATGTCCTTTGAGGACAGCATGTGGAGA 2822
Db 2760 TGCACACACAGTGTCTAAGCAGTCCGACACCATGTCCTTTGAGGACAGCATGTGGAGA 2819
QY 2823 TTGCACACAGGCGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTC 2882
Db 2820 TTGCACACAGGCGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTC 2879
QY 2883 CAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2942
Db 2880 CAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2939
QY 2943 CTGCCCCCTGAAATTTGTCAGGCTACTGCTACCTGTAGTCAATGCTTTGGAGCAACAGG 3002
Db 2940 CTGCCCCCTGAAATTTGTCAGGCTACTGCTACCTGTAGTCAATGCTTTGGAGCAACAGG 2999
QY 3003 CTGTGGCTGTGTACTGATCCAGCAATCTGCAAGGGAATGCAATAGAGGTTCTCTA 3062
Db 3000 CTGTGGCTGTGTACTGATCCAGCAATCTGCAAGGGAATGCAATAGAGGTTCTCTA 3059
QY 3063 TAAAGACACAGTGAAGATGCTTCCGACAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122
Db 3060 TAAAGACACAGTGAAGATGCTTCCGACAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119
QY 3123 GCTCAATTCACAGATGTCTAGAGACAGGAGATACAACTGGTCTTTTCACTCACTGCC 3182
Db 3120 GCTCAATTCACAGATGTCTAGAGACAGGAGATACAACTGGTCTTTTCACTCACTGCC 3179
QY 3183 AGCTTGCCCAATCAAGGCGCAGTAATGCAATCAATCAGACATCTGTGAGAAAGTGTGA 3242
Db 3180 AGCTTGCCCAATCAAGGCGCAGTAATGCAATCAATCAGACATCTGTGAGAAAGTGTGA 3239
```

```
QY 3243 GAACCTGACACAGGCAAGCACTGGAGACCTGTCATATCTGGCTTCTACGGTGTATCCAC 3302
Db 3240 GAACCTGACACAGGCAAGCACTGGAGACCTGTCATATCTGGCTTCTACGGTGTATCCAC 3299
QY 3303 CAATGAGGGAATGCTCAGCCATGCAAGTGCATGCAATGGGCACCGCTCTCTGTGCAACACCAA 3362
Db 3300 CAATGAGGGAATGCTCAGCCATGCAAGTGCATGCAATGGGCACCGCTCTCTGTGCAACACCAA 3359
QY 3363 CACGGGCAAGTGTCTTCGACACCAAGGGGCTCAGGGGACAGTCCAGCTATGTGA 3422
Db 3360 CACGGGCAAGTGTCTTCGACACCAAGGGGCTCAGGGGACAGTCCAGCTATGTGA 3419
QY 3423 GGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3482
Db 3420 GGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3479
QY 3483 TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATCGCTATTAATTCAGAGCTATCAA 3542
Db 3480 TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATCGCTATTAATTCAGAGCTATCAA 3539
QY 3543 TTTTGTGGTACTCTCTGACGACAAACAGGATTTGGACATGTTTCATCAATGCCCTCCAA 3602
Db 3540 TTTTGTGGTACTCTCTGACGACAAACAGGATTTGGACATGTTTCATCAATGCCCTCCAA 3599
QY 3603 GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGA 3662
Db 3600 GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGA 3659
QY 3663 AGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAAATGA 3722
Db 3660 AGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAAATGA 3719
QY 3723 GAAGTTTGATTTTTCGCAACCCCAATATCACTTTCTTTGTTTATGTGAGTAATTTTCA 3782
Db 3720 GAAGTTTGATTTTTCGCAACCCCAATATCACTTTCTTTGTTTATGTGAGTAATTTTCA 3779
QY 3783 CTGGCCCATCAAAATTCAGAT 3803
Db 3780 CTGGCCCATCAAAATTCAGAT 3800
```

RESULT 3

```
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-1
```

```
Query Match 79.2%; Score 3398; DB 3; Length 8827;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 3869; Conservative 0; Mismatches 410; Indels 205; Gaps 2;
QY 3 GGTGGCGCAGCGCGCGCAACTGAGGCAAGCTGAGAGAGGAGCGCGCGCAGCGCAGC 62
Db 82 GGTGGCGCAGCGCGCGCGCACTGAGGCGCGCTGAGGCGGAGCAGGAGCAGCAGC 141
```


Db	2293	CCTCTGGGAGTTGGCAACTGAAGAAACAAGCAAGAAAGTTAAATCAGAGTGTCTTTCTTA	2359
Qy	2087	AAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATTGTTACAGCTGTACAGCCA	2146
Db	2353	AAAGAACCTTTGACCATGACAGATGTGACAGCACACAGATTGTTACAGCTGCACAGCCA	2412
Qy	2147	ACACCAATGACTGCCACTGGTGCAATGACCATTTGTCTCCAGGAGACACACAGCTGCTCAG	2206
Db	2413	ATAACCAATGACTGCCACTGGTGCAATGATCATTGTCTCTCTGTGAACCAACAGCTGCACAG	2472
Qy	2207	AAGGCCAGATCTCCATTTTATAGTATGAGAAATGGCCCCAAGGATAACCCCATGTACTACT	2266
Db	2473	AAGGCCAGATCTCCATTTGCCAAGTATGAGAGTTGCCCCAAAGGATAACCCCATGTACTACT	2532
Qy	2267	GTAAACAAGAGAACAAGTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCC	2326
Db	2533	GCAATGAAGAAACAGCTGCAGAGCTGTGCCCTAGACAGAACTGCCAGTGGAGCCCC	2592
Qy	2327	GGAAATCAGGAGTGCATTTGCCCTGCCCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTG	2386
Db	2593	GGAAATCAGAGTGCATGCCCTGCCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTG	2652
Qy	2387	GAACCTCATGTTTGAATAATTAATCTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCT	2446
Db	2653	GAACCTCGTGTGAAATCACTACTGCTAAGGAGAAATATGACAAATGCTAAATTTGTTCT	2712
Qy	2447	GTAGGAACCAATGCCCTTTTGGCTTCTCTTACAACCCAGAGAAGGTAGAAATTTGTCC	2506
Db	2713	GTAGGAACCAATGCCCTTTTGGCTTCTCTCTCAATCCAGAGAAGGTGGAGTTTGTCC	2772
Qy	2507	TTAAGCAGCTGCCAATAATCAGTCAATCTCAGAGCATGTCCAAGCTCACTTAACCCCAT	2566
Db	2773	TTAAGCAGCTTCGATTAATCAATCACTCAAAAGTATGTCCAAGCTCACTGTGACTCCAT	2832
Qy	2567	GGGTGGCCCTTCGAAACATCAATGTGCTCTACTGGTCTGGGAAGATATGTCGCCATTTA	2626
Db	2833	GGGTGTGCTTCGGAAGATCAATGTGCTTACTGGTCTGGGAGGATATGCTTCCATTCA	2892
Qy	2627	CAAAATGTTTACTACAGTGCATGCCGTCTCAGAGCCAGTCAATGTGGATTCTGTGGAATTT	2686
Db	2893	CAAAATGTTTGTCTGCACTGGATGCCATCTCAGAGCCAGTCAATGTGGGATCT	2952
Qy	2687	TATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTCGCATCAACCCACTCAATGGTA	2746
Db	2953	TGTCAAGAGCCTTAGTACTCGGGGATTAAGGCTGCAACCTCGCATCAACCCCTCAATGGCA	3012
Qy	2747	GTGTCTGTAAGGCTTCGAAAACACACAGTGTCAAGCAGTCCGACACCAATGTGCTCTTGA	2806
Db	3013	CGGTCTGTGTAAGGCTTCGAAAACACAGTGTCAAGCAGTCCGACACCAATGTGCTCTTGA	3072
Qy	2807	GGACAGCATGTGGAGATTGCAACCAAGCGCAGCTCTGAGTCATGTGGTCAGCAACAATGA	2866
Db	3073	GGACAGCGTGTGGCGAGTGCATAGCAGCAGCTCCGAGTGCATGTGGTCAGTAAATGA	3132
Qy	2867	AGCAGTGTGTCGACTCCAAATGCCCTATGTGGCTCCCTTCCCTTTGGCCAGTGTATGAAAT	2926
Db	3133	AGCAGTGTGTCGACTCCAAATGCCCTATGTGGCTCCCTTCCCTTTGGCCAGTGTATGAAAT	3192
Qy	2927	GGTATACCATCAGCACCTGCCCTCCCTGAAAATTGTTCCAGGCTACTGTACCTGTAGTCAAT	2986
Db	3193	GGTATACCATCAGCAGCTGCCCTCCCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATT	3252
Qy	2987	GCTTGGAGCAACCAAGCTGTGGCTGTGTTACTGATCCAGCAATACTTGGCAAGGGAAT	3046
Db	3253	GCTTGGAGCAGCAAGCTGTGGTGTGTTACTGATCTCTAGCAATACTTGGCAAGGGAAT	3312
Qy	3047	GCATAGAGGGTCTCTATAAAGACCAAGTGAAGATGCCCTTCCAGGCCCTTACAGGAAAT	3106
Db	3313	GTATTTAGGGCAGCTATATAAGGACCTGTGAAGATGCCCTCAGAGCCTCTCGGAAATG	3372
Qy	3107	TCTATCCACAGCCCTGCTCAATTCACGATGTGTGTAGAGGACAGCAGATACAACTGGT	3166
Db	3373	TGTATCCACAGCCCTTCTGAACTCCAGCATGTGTGTAGAGGACAGCAGATACAACTGGT	3432

Qy	3167	CTTTTCATTCTCACTGTCTCAGCTTGGCAATGCAACGGCCACAGTAATATGCATCAATCAAGCA	3222
Db	3433	CTTTTCATTCTCACTGTCTCAGCTTGGCAATGCAACGGCCACAGTAATATGCATCAATCAAGCA	3492
Qy	3227	TCTGTGAGAAGTGTGAGAACCTGCACACAGGCAAGCACTGCGAGACCTGCATATCTGGCT	3286
Db	3493	TCTGTGAGAAGTGTGAGAACCTGCACACGGGCAAGCACTGCGAGACCTGCATATCTGGCT	3552
Qy	3287	TCTACGGTGATCCACCAATGGAGGGAATGTGCGCCATGCAAGTGCATATGGGCAAGCGCT	3346
Db	3553	TCTATGGTGACCGCATTAATGGAGGCAAAATGTGCGCCATGCAAGTGCATATGGGCAAGCGCT	3612
Qy	3347	CTCTGTGCACCAACCAACACGGGCAAGTGTCTGCAACCAACCAAGGCGCTCAAGGGGAGCG	3406
Db	3613	CACCTGTGCACCAACCAACACGGGCAAGTGTCTGTAACCAACCAAGGCTCAAGGGGAGCG	3672
Qy	3407	AGTGCCAGCTATGTGAGGTAGAAATTCGATACCAAGGAAACCTCTCAGAGGAACAATGTT	3466
Db	3673	AGTGCCAGCTATGTGAGGTAGAAATTCGATACCAAGGAAACCTCTCAGAGGAACAATGTT	3732
Qy	3467	ATTATATCTCTTTATTTGATCATCAGTTCACTTTTAGTCTTATCCAGGAAGATGATCGCT	3526
Db	3733	ACTATACCTTTCTCAATGACTATCATCAGTTCACTTTTAGCTTGTCCAGGAAGACGACCGCT	3792
Qy	3527	ATTACACAGCTATCAATTTTGTGGCTACTCTGACGAAACAAACACAGGATTTTCGACATGT	3586
Db	3793	ACTACACAGCCATCAACTTTGTGGCTACTCTGATGAAACAAACACAGGATTTTCGACATGT	3852
Qy	3587	TCATCAATGCTCCAGAAATTTCAACCTCAACATCACTGGGTGCAGTTTCTACGCTG	3646
Db	3853	TCATCAATGCTCCAGAAATTTCAACCTCAACATCACTGGGTGCAGTTTCTCCAGCCG	3912
Qy	3647	GAACCCAGGCTGGAGAGAGATGCGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAG	3706
Db	3913	GAACCCAGATGGAGAAAGAGTGCCTGTGTTTCAAAAACCAACATCAAGGATACAAG	3972
Qy	3707	ATAGTTCTCTTAATGAGAAGTTTGATTTTCCCAACCAACCAATATCACTTCTTTGTTT	3766
Db	3973	ATAGTTCTCTTAATGAGAAATTTGATTTTGGCAACCAATCAACATCACTTCTTTGTTT	4032
Qy	3767	ATGTCAGTAATTCACCTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGACAGAAAT	3826
Db	4033	ATGTCAGTAATTCACCTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGACAGAAAT	4092
Qy	3827	TTATGGACCTGGTACAGTTCTTCGTGACTTCTTCAGTGTGTTTTCCTCTCTTTGCTCTGG	3886
Db	4093	TCATGGACCTGGTACAGTTCTTCGTGACTTCTTCAGTGTGTTTTCCTCTCTTTGCTCTGG	4152
Qy	3887	TGGCTGCTGGTTTGGAGATCAACAAAGTTGTTTGGGCGCTCCAGACGTAGAGAGCAAC	3946
Db	4153	TGGCTGAGTGGTCTGGAAGATCAAGCAGAGCTGTGGGGCATCCAGCGGAGAGAGCAAC	4212
Qy	3947	TTCTTCGAGAGATCAACAGATGCCAGCCGTCCCTTTGCTCTGTATAATGTGCTCTTGG	4006
Db	4213	TTCTTCGGGAGATCAACAGATGCCAGCCGTCCCTTTGCTCTGTATAATGTGCTCTTGG	4272
Qy	4007	AAAACAGATGAGAGCGCTCTGATCTTAATGGGGGGAGTAAAGACATGTTCCCAACCCCA	4066
Db	4273	AAAACAGATGAGAGCGCTCTGATCTTAATGGGGGGAGTAAAGACATGTTCCCAACCCCA	4332
Qy	4067	TTGCACTGGAGCCGTGTTTGGCAACAAAGCCGTGCTCTCTGTGTGTTTGTGAGGCTCC	4126
Db	4333	TTGCACTGGAGCCGTGTTTGGCAACAAAGCCGTGCTCTCTGTGTGTTTGTGAGGCTCC	4392
Qy	4127	CTCGAGGCTGGGTGGATCCCTCTCTCTGGGCACTCAGGTCTTGTGTGGCCACAGCCCTC	4186
Db	4393	CTCGAGGCTGGGTGGATCCCTCTCTCTGGGCACTCAGGTCTTGTGTGGCCACAGCCCTC	4452
Qy	4187	TGTTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAGTCAAGAGCCGTGAGAA	4246
Db	4453	TGTTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAGTCAAGAGCCGTGAGAA	4512

Db 1580 AGTGGTGGACCAATGCTGGTGGTGGGGAACACACACAAATGACACATCTATGACCA 1739
Qy 1743 TGGGCGCAATGCTTCTTTCAGATTCATGGCCCTATGACATTCGCTGTGACCGCTGGTC 1802
Db 1740 TGGGCGCAATGCTTCTTTCAGATTCATGGCCCTATGACATTCGCTGTGACCGCTGGTC 1799
Qy 1803 AGTCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTCTT 1862
Db 1800 AGTCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTCTT 1859
Qy 1863 ACACACAGCACCATGATGTTGGTGGTGGTTCATATGATCTCTCTCAGCGACATCTCT 1922
Db 1860 ACACACAGCACCATGATGTTGGTGGTGGTTCATATGATCTCTCTCAGCGACATCTCT 1919
Qy 1923 GGTATTCACCTCGGACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTATGACGAGG 1982
Db 1920 GGTATTCACCTCGGACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTATGACGAGG 1979
Qy 1983 ACCTGGTATTCGGTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTCTGGGCGCTGGC 2042
Db 1980 ACCTGGTATTCGGTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTCTGGGCGCTGGC 2039
Qy 2043 AACTGATGAACAAGAAAGTTAAATCAGATGTTTTCGAAAGAACTCTTGACCA 2102
Db 2040 AACTGATGAACAAGAAAGTTAAATCAGATGTTTTCGAAAGAACTCTTGACCA 2099
Qy 2103 TGACAGATGTACACAGACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCA 2162
Db 2100 TGACAGATGTACACAGACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCA 2159
Qy 2163 CTGCTGCAATGACCATGTTGTCGCCAGAACCAACAGCTGTCTCAGAGCCGACATCTCCAT 2222
Db 2160 CTGCTGCAATGACCATGTTGTCGCCAGAACCAACAGCTGTCTCAGAGCCGACATCTCCAT 2219
Qy 2223 TTTTAGGTATGAGATTCGCCAGGATTAACCCATGATCTACTGTAACAGAGACAG 2282
Db 2220 TTTTAGGTATGAGATTCGCCAGGATTAACCCATGATCTACTGTAACAGAGACAG 2279
Qy 2283 CTGACGAGCTGTGCTCTGGACCAAGCTGCGAGTGGAGCCCGGAATCAGGAGTGCAT 2342
Db 2280 CTGACGAGCTGTGCTCTGGACCAAGCTGCGAGTGGAGCCCGGAATCAGGAGTGCAT 2339
Qy 2343 TGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGG 2387
Db 2340 TGCCCTGCCGAGTAGGCTCTGACGGTCTATCTTGGTGTGTGGG 2384

RESULT 5

US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-12

Query Match

43.2%; Score 1853; DB 3; Length 6370;

Best Local Similarity 96.3%; Pred. No. 0;
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;
Qy 2257 ATGTACTACTTAAACAAGNAGACGAGCTGCAGAGCTGTGCCCTGAGACAGAACTGCCAG 2316
Db 1 ATGTACTACTTAAACAAGNAGACGAGCTGCAGAGCTGTGCCCTGAGACAGAACTGCCAG 60
Qy 2317 TGGAGCCCGGGAATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCAATTGGCTGG 2376
Db 61 TGGAGCCCGGGAATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCAATTGGCTGG 120
Qy 2377 CATTTGGTTGGAATCTCATGTTTGAATATTACTCTGCCAAGGAGAAATATGACAACTGT 2436
Db 121 CATTTGGTTGGAATCTCATGTTTGAATATTACTCTGCCAAGGAGAAATATGACAACTGT 180
Qy 2437 AAATTTCTCTGTAGGAACCAACAATGCCCTTTTGGCTTCTCTTACAAACCCAGAAAGGTA 2496
Db 181 AAATTTCTCTGTAGGAACCAACAATGCCCTTTTGGCTTCTCTTACAAACCCAGAAAGGTA 240
Qy 2497 GAATTTGCTTCTTAAGCAGCTGCGAATATATGAGTGCATCTCAGAGCATGTCCAAAGCTCAC 2556
Db 241 GAATTTGCTTCTTAAGCAGCTGCGAATATATGAGTGCATCTCAGAGCATGTCCAAAGCTCAC 300
Qy 2557 TTAACCCCATGGTGGGCTT---CGGAAGATCAATGTCTCTACTGG---TGCTGGG 2608
Db 301 TTAACCCCATGGTGGGCTT---CGGAAGATCAATGTCTCTACTGG---TGCTGGG 360
Qy 2609 AAGATATGTCCTTACAAATAG-TTTPACTACAGTGGATGCC---GTCTGAGGCCAG 2663
Db 361 GATATGTCCTTACAAATAGTTTACTACAGTGGATGSCCGTCTTGAGGCCAG 420
Qy 2664 TGAATGCT--GGATTTCTGTGGAATTTT---ATCAGAACCCAGTACTCGGAGCTGAAAG 2716
Db 421 TGTGCTTGGATTTCTGTGGGAATTTTATTCAGGAACCCAGTTTACTTTGGGGACTGAAAG 480
Qy 2717 CTGCAACCTGTCAT--CAACCCCATCAATGATGATGTCTGTGAAAGCTCTGCAACCCAG 2774
Db 481 CTGCAACCTGTCAT--CAACCCCATCAATGATGATGTCTGTGAAAGCTCTGCAACCCAG 540
Qy 2775 TGCTAA-GCAGTGCAGGACCAACATGTCCTTGAAGCAGCATGTGGAGATTCGACCCAGG 2833
Db 541 TGCTAAGGAGTGCAGGACCAACATGTCCTTGAAGCAGCATGTGGAGATTCGACCCAGG 600
Qy 2834 GCAGCTCTGAGTGCATGTGTGAGCAGCAACATGAAGCAGTGTGTGGAATCCATTCGCTATG 2893
Db 601 GCAGCTCTGAGTGCATGTGTGAGCAGCAACATGAAGCAGTGTGTGGAATCCATTCGCTATG 660
Qy 2894 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATGATGAGCAGTGTGAGCAGTGTGAGCAG 2953
Db 661 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATGATGAGCAGTGTGAGCAGTGTGAGCAG 720
Qy 2954 AAAATTTGTTGAGCTACTGTACTGTAGTCAATGTTGGAGCAACCCAGCTGTGGCTGTGT 3013
Db 721 AAAATTTGTTGAGCTACTGTACTGTAGTCAATGTTGGAGCAACCCAGCTGTGGCTGTGT 780
Qy 3014 GTACTGATCCAGCAATATCTGCAAGGGAATGCAATAGAGGTTCTCTATAAGGACCAG 3073
Db 781 GTACTGATCCAGCAATATCTGCAAGGGAATGCAATAGAGGTTCTCTATAAGGACCAG 840
Qy 3074 TGAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCTGCTCAATTCGA 3133
Db 841 TGAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCTGCTCAATTCGA 900
Qy 3134 GCATGTGTCTAGAGGACAGCAGATCAACTGTTGTTTCTTCTACTGTCCAGCTTGGCAAT 3193
Db 901 GCATGTGTCTAGAGGACAGCAGATCAACTGTTGTTTCTTCTACTGTCCAGCTTGGCAAT 960
Qy 3194 GCAACCGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCA 3253
Db 961 GCAACCGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCA 1020
Qy 3254 CAGCAGGACACTGGAGACCTGCATATCTGCTTCTACGGTGTATCCAGCTGAGGAGGA 3313

```
Db 1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGAGGGA 1080
QY 3314 AATGTCAGCCATGCAAGTCAATGGCCAGCGCTCTGTGCAACCAACACGCGCAAGT 3373
Db 1081 AATGTCAGCCATGCAAGTCAATGGCCAGCGCTCTGTGCAACCAACACGCGCAAGT 1140
QY 3374 GCTTCTGCAACCAACGAGCGCTCAAGGGGAGCGAGTGCAGCATGTGAGGTAGAAAATC 3433
Db 1141 GCTTCTGCAACCAACGAGCGCTCAAGGGGAGCGAGTGCAGCATGTGAGGTAGAAAATC 1200
QY 3434 GATACCAAGAAACCTCTCAGAGAAACATGTTATATATCTTCTTATATGCACTATCAGT 3493
Db 1201 GATACCAAGAAACCTCTCAGAGAAACATGTTATATATCTTCTTATATGCACTATCAGT 1260
QY 3494 TCACCTTTAGTCTATCCCAAGGAAGATGCGCTATTACACAGCATATCAATTTTGGGCTA 3553
Db 1261 TCACCTTTAGTCTATCCCAAGGAAGATGCGCTATTACACAGCATATCAATTTTGGGCTA 1320
QY 3554 CTCCTGACGAAACAAACAGGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAAC 3613
Db 1321 CTCCTGACGAAACAAACAGGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAAC 1380
QY 3614 TCAACATACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCGTG 3673
Db 1381 TCAACATACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCGTG 1440
QY 3674 TTGTTTCAAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATCAGAAAGTTTGATT 3733
Db 1441 TTGTTTCAAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATCAGAAAGTTTGATT 1500
QY 3734 TTCGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTACCTGGCCATCA 3793
Db 1501 TTCGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTACCTGGCCATCA 1560
QY 3794 AAATTCAGATTGCCCTTCTCTCAGACACAGCAATTTTATGGAAGCTGGTACAGTTTCGTGA 3853
Db 1561 AAATTCAGATTGCCCTTCTCTCAGACACAGCAATTTTATGGAAGCTGGTACAGTTTCGTGA 1620
QY 3854 CTTTCTCAGTTGTTTCTCTCTTTGCTCTGTTGGCTGCTGTGGTTTGAAGATCAAAAC 3913
Db 1621 CTTTCTCAGTTGTTTCTCTCTTTGCTCTGTTGGCTGCTGTGGTTTGAAGATCAAAAC 1680
QY 3914 AAAGTTGTTGGCTTCCAGAGCTAGAGAGCAACTTCTTCAGAGATGCAACAGATGGCCA 3973
Db 1681 AAAGTTGTTGGCTTCCAGAGCTAGAGAGCAACTTCTTCAGAGATGCAACAGATGGCCA 1740
QY 3974 GCGTCCCTTTGCCCTCTGTAATGTCGCTTTGGAACAGATGAGAGGCTCCTGATCTTA 4033
Db 1741 GCGTCCCTTTGCCCTCTGTAATGTCGCTTTGGAACAGATGAGAGGCTCCTGATCTTA 1800
QY 4034 TTGGGGAGTATAAGACTGTTCCCAACCCATTCACCTGGAGCCGTGTTTGGCAACA 4093
Db 1801 TTGGGGAGTATAAGACTGTTCCCAACCCATTCACCTGGAGCCGTGTTTGGCAACA 1860
QY 4094 AAGCGCTGTCTCTCTCTGTTTGTGAGGCTCCCTCGAGGCTCGGTTGGCATCCCTCCTC 4153
Db 1861 AAGCGCTGTCTCTCTCTGTTTGTGAGGCTCCCTCGAGGCTCGGTTGGCATCCCTCCTC 1920
QY 4154 CTGGCAGTCAAGTCTTCTCTGTTGAGGCTCGGCTGGTGGACATTTCTCAGCAGATGCCGA 4213
Db 1921 CTGGCAGTCAAGTCTTCTCTGTTGAGGCTCGGCTGGTGGACATTTCTCAGCAGATGCCGA 1980
QY 4214 TAGTGTACAAGGAGAGTCAAGAGCGGTGAGAAACCGAAGCAGCAGCCCTGSCAGC 4273
Db 1981 TAGTGTACAAGGAGAGTCAAGAGCGGTGAGAAACCGAAGCAGCAGCCCTGSCAGC 2040
QY 4274 CTGGACCTGCATCTGA 4290
Db 2041 CTGGACCTGCATCTGA 2057
```

```
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-245-041-8
```

```
Query Match 35.5%; Score 1521.8; DB 3; Length 2419;
Best Local Similarity: 80.7%; Pred. No. 0;
Matches 1895; Conservative 0; Mismatches 247; Indels 205; Gaps 2;

QY 3 GGTGGCGGCGAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGGCGAGC 62
Db 82 GGTGGCGGCGGCGGCGGCGACTGAGGCGGCGCTGAGGGGAGGAGGAGGAGCGGCGGCGAGC 141
QY 63 GCTGGCGGCGAGAGCGGCGGCGGCGACTGAGGAGCTGGGAGCTGAGGAGCTGAGGAGGCGG 122
Db 142 GCTGGCGGCGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 201
QY 123 GGGGCTGGGGGCGGCGGCTGCGCTCCCGGCGGCTGCTCCACCGCTGCGGCGGCGGCGGCT 182
Db 202 GGGAGCGGCGGCGGCGGCTGCTGCTCCCGGCGGCTGCTGCTGGGGGCGGCTGCGGCGGCG 261
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
Db 262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
QY 243 CGAGGCGGCGGCGGCGGCGGCGGCGGCTGCTGCGGCTGCTGCGGCGGCGGCGGCGGCGGCG 302
Db 313 CGAGGCGGCTGCGGCTGCGGCGGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG 372
QY 303 TGACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362
Db 373 TGACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
QY 363 CGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 410
Db 433 CAGCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 492
QY 411 ----- 410
Db 493 ACGCCTATATACAGCTGTTTCGAGAGGTGAGGCTGGAGGAGACAGTTTCGAGGCAAGCTTC 552
QY 411 ----- 410
Db 553 GGCTACAGAAATAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTTCCAAACCAAAA 612
QY 411 ----- 410
Db 613 TGAGCGAAAGAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTGACTAATTTTTCACC 672
QY 411 ----- ACTAATCGATCTTCTGGGTTTGTGACAGATGAGACCTGGAAATTTATAATACAAA 466
Db 673 GAGAACTAACTGGCTCTTCTGGATTGTTAAACAGATGGACCTGGGAAATTTATAATATAAG 732
QY 467 CGAAGTGCAGCGGCTCATTTGAAGACAGCAAGCAATAGATATAGATGAGCTTCTGTTTCATC 526
Db 733 CGAAGTGCAGATGGCTCATTTGAAGACAGCAAGCAATAGATATAGATGAGCTTCTGTTTCACC 792
```

QY	527	ATTTTGCTACAGAGTGTAGTTGGACCATTATATATGTTTATGATGGGACTCAATTTATG	586
DB	793	ATTTTGCTACAGAATGTAGCTGGGACCAATTATGTTTATGATGGGACTCAATCTACG	852
QY	587	CACCGTAGTGGCTGCATTTAGTGGCCTCATTTCTCTGAGAGAGATGCCAATGAGACTG	646
DB	853	CACCTCGATTGCTGCTTTAGTGGCCTCATTTCTCTGAAGAGATGCCAATGAGACCG	912
QY	647	TCGCTGAGGTTGTGGCACATCAGGTTATGCCCTTGTCGCAATTTTTTAGTGATGCTGCTT	706
DB	913	CTCTGAGGTCACCTGTCACTTCAGGTTATGCACTGCTGCAATTTTTTAGTGATGCTGCTT	972
QY	707	ATAATTTGACTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	766
DB	973	ATAATCTGACTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1032
QY	767	GCCGAGGAGAGTGTAAAGATCAGTAATAGCAGCGAAACTGTGTGAATGTGAATGTTCGAAA	826
DB	1033	GCCGAGGAGAGTGTAAAGAGCAGTAACAGCAGCAGCGCTGTGTGAATGTTCGAAA	1092
QY	827	ACTGGAAGAGTGAAGCATGTGACATTCCTCACTGTACAGACAACACTGGGTTTTCCCTCANC	886
DB	1093	ACTGGAAGAGGAGTGTGTGACATTCCTCACTGTACAGACAACACTGGGTTTTCCCTCACC	1152
QY	887	GAGGCATCTGCAATCCAAGTCATGTACAGAGATGCTCCTGCTTCTCAGACTGCCAGGGTC	946
DB	1153	GAGGCATCTGTAATGCAAGCGATACAGAGGGTCTCCTGCTTCTCAGTGGCAGGGTC	1212
QY	947	CTGGAATGTTCAAGTTCCTGTACCAGCTAACAGTCAATTTTGGACTCGAGAGGATATTTCTA	1006
DB	1213	CTGGAATGTTCAATTCCTGTGCCAGCTAACAGTCAATTTTGGACTCGAGAGGATATTTCTG	1272
QY	1007	ACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACAATTTATGTGGTTG	1066
DB	1273	ATTTAAAGCTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAACAATTAATGTGGTTG	1332
QY	1067	TTGAGAGATATATGTTTCAAACCATCAGATTAATTAACATGGTCTTAGCGTATGACCTTGCTT	1126
DB	1333	TTGCGGATATATGTTTCAAACCATTCAGTTACAGCATGGTCTTAGCGTATGACCTGACCT	1392
QY	1127	CTAGGAGTGGCTTCCACTAAACCGTTCTGTGAAACAATGTGGTTGTTAGATATGGTCAAT	1186
DB	1393	CTAGGGAATGGCTTCCACTAAACCATCTGTGAAACAGTGTGGTTGTTAGATATGGTCAAT	1452
QY	1187	CTTTGGCATTTACAGGATTAATTTTACATGTAATGAGGAAAAAATGATCCAACTGGGA	1246
DB	1453	CTTTGGCATTTACATGAGGATTAATTTTACATGTAATGAGGAAAAAATGATCCAACTGGGA	1512
QY	1247	ATGTGACCAATGAGTTGAGAGTCTTTCACATTCATTAATGAGTCAATGGGTGTGTTGACCC	1306
DB	1513	ACGTGACCAATGAGCTGAGATTTTTCATATTCATTAATGAGTCAATGGGTATTTGTTAACTC	1572
QY	1307	CTAAGGCAAGAGGAGCATGATCGAGTGTGGGCACTCTGACACATTTGTTTACACTGAAGA	1366
DB	1573	CGAAAGCTTAAGGATCAGTATGACGTGGTGTGACACTCAGCACATTTGTTACACTGGCAT	1632
QY	1367	ATGGCCGAGTGTGATGCTGTGTCATCTTTGGTCTACTGCCCTCTCTATGGATATATAAGCA	1426
DB	1633	CTGGCCGTGTGGTCAATGTTGGTCACTCTTCGGTCAATGCCCCATCTATGGATATATAAGC	1692
QY	1427	ATGTGACGAATATAGATTTGGATTAAGAACACATGGAGTATATTTACACACCCAGGGTGCC	1486
DB	1693	TGTGTCAGGAATATGACTTGGAAAGAACACATGGAGTATATTTACACTCAGGGTGCTC	1752
QY	1487	TGTGTCAGAGGGGTTTACGGCCATAGCAGTGTTTAGCAACATPAGACACAGGGGCCCTATACG	1546
DB	1753	TGTGTCAGAGGGGTTTATGGCCACAGTAGTGTTTATGTGACAGGACCAAGGCTCTGTACG	1812
QY	1547	TTTCATGGTGGCTACAAGGCTTTTCAAGTGCATTAAGTACCGGCTTGCAGATGATCTTACC	1606
DB	1813	TTTCATGGTGGCTACAAGGCTTTTCAAGGCTTTCAGGCAACAAATACCGGCTTGCAGATGATCTTACA	1872

```

RESULT 7
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: fastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10

```

		Query Match	14.9%	Score	540.2;	DB	3;	Length	1051;
		Best Local Similarity	86.1%;	Pred.	No.1.4e-154;				
		Matches	724;	Conservative	0;	Mismatches	106;	Indels	9;
		Gaps							17;
Qy	3	GTTGGCCGACGGCGGCCCACTAGGCMAAGCCTAGAGGAGGAGCGCGGCGGACGGCAGC	62						
Dd	216	GTTGGCGCGCGCGCGCGACTAGAGGCGCGCCTAGGGGGGAGCAGCAGCACATGACGAC	275						

Db 186 CTTCCGAAGATCAATGTCTCTACTGTGTGCTGGAGAGATATGCCCCATTACAAATAGT 127
Qy 2635 TTACTACAGTGGATGCGCTGTGAGCCAGTGATGCTGGATTCGTGGAAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGATGCTGGATTCGTGGAAATTTTATCAGAA 67
Qy 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 7
Qy 2755 GAAAGG 2760
Db 6 GAAAGG 1
RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966
Query Match 5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.4e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 187
Qy 2575 CTTCCGAAGATCAATGTCTCTACTGTGTGCTGGAGAGATATGCCCCATTACAAATAGT 2634
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGTGCTGGAGAGATATGCCCCATTACAAATAGT 127
Qy 2635 TTACTACAGTGGATGCGCTGTGAGCCAGTGATGCTGGATTCGTGGAAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGATGCTGGATTCGTGGAAATTTTATCAGAA 67
Qy 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 7
Qy 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966
Query Match 5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.4e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 187
Qy 2575 CTTCCGAAGATCAATGTCTCTACTGTGTGCTGGAGAGATATGCCCCATTACAAATAGT 2634
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGTGCTGGAGAGATATGCCCCATTACAAATAGT 127
Qy 2635 TTACTACAGTGGATGCGCTGTGAGCCAGTGATGCTGGATTCGTGGAAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGATGCTGGATTCGTGGAAATTTTATCAGAA 67
Qy 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 7
Qy 2755 GAAAGG 2760
Db 6 GAAAGG 1
RESULT 12
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833.381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516.448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1916
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(415)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916
Query Match 5.0%; Score 212.8; DB 4; Length 415;
Best Local Similarity 73.6%; Pred. No. 6.5e-45;
Matches 271; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 422 CTTCTGGGTTTGACAGATGACCTGGAAATTAATAACAAAGAGTGCAGTGGC 481
Db 5 CCGTCGGATATTAAACAGATGGCCCAATTAACATATAATAAATAAACTAAATGTAATTGGC 64
QY 482 TCATTGAAGCAGACGCAATAGAAATAGAGACTTCGTTTCAATCAATTTTCTCAGAGT 541
Db 65 TCATTGAAGCTATCCAAATGAGTGTAAAGATTCAATTTTCTCAGAGT 124
QY 542 GTAGTTGGACCAATTAATGTTTATGATGGGAGTCAATTTATGACCCGCTAGTGTG 601
Db 125 GTAGCTGGGATCATATGATGTTTATGATGGGAGTCAATTTATGACCCGCTAGTGTG 184
QY 602 CATTAGTGGCTCATTGTTCTGAGAGAGATGGCAATGAGACTGCTCCTCAGGTTGTTG 661
Db 185 TACTTAGTGGTTGATAGTCCCTGAAATAGGGGCATGAACTGTGCTGAAATGTTGTA 244
QY 662 COACATCAGGTTATCCCTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGAT 721
Db 245 CTACATCTGGCTATGCACTGTATCAATTTTATGATGCTGCTGCTTATTAATTTGACT 304
QY 722 TTAATATTACTTTACAGTTTGTATGATGTTCCAAATGAGTCTCAGGCGGAGAGTGA 781
Db 305 TCAACATTTCTTATCAATCAATCTTGTCTTAACATGCTCTGGTCAATGGGAAGTGA 364
QY 782 AGATCAGT 789
Db 365 CAACATGAT 372

RESULT 13
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

Query Match 4.7%; Score 201.2; DB 3; Length 90050;
Best Local Similarity 81.5%; Pred. No. 1.5e-40;
Matches 233; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1169 TTGTTAGATATGGCTATCTTTGGCATTATCAAGGATAAAATTTACATGATGGAGAA 1228
Db 6931 TGGTTTAAAAATCCTGTTTACATCTTCTTATAGGATAAAATCTACATGATGGAGAA 6990
QY 1229 AAATGATCCAACTGGGAATGTGACCAATGATGAGAGTTTTCATTTCAATTAATGAGT 1288
Db 6991 AAATGATTTCAAGGAGCACTGACCAATGATGAGAGTTTTCATTTCAATTAATGAGT 7050
QY 1289 CATGGTGTGTTTGAACCCCTAAGCAAGGAGCAGTATGAGTGGTGGTGGCCTCTGAC 1348
Db 7051 CATGGTATGTTTAACTCCGAAAGCTAAGGATCAGTATGAGTGGTGGTGGCCTCTGAC 7110
QY 1349 ACATGTTACACTGAAGAGTGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1408
Db 7111 ACATGTTACACTGGCATCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7170

QY 1409 TCTATGATATATAAGCAATGTGACGAATATGATTTGGATTAAGAA 1454
Db 7171 TCTATGATATATAAGCGTTGTGACGAATATGATTTGGTATGTA 7216
RESULT 14
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-3

Query Match 3.9%; Score 169.2; DB 3; Length 17056;
Best Local Similarity 86.9%; Pred. No. 9.5e-33;
Matches 186; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 732 TTACAGTTTGTATGTGTCCTCAATTAACCTGCTCAGGCGGAGAGTCTAAGATCACTAA 791
Db 4090 TTGTAGTTTGTACATGTGTCCTCAATTAACCTGCTCAGGCGGAGAGTCTAAGATCACTAA 4031
QY 792 TAGCAGCGAAACTGTTGATGTAATGTAATGTTCTGAAACTGGAAGGATGAGCATGTGACAT 851
Db 4030 CAGCAGCAGCGCTGTTGATGTAATGTTCTGAAACTGGAAGGATGAGCATGTGACAT 3971
QY 852 TCTCTACTGTACAGCAACTGTGGTTTCTCTCATCGAGCATCTCCAATTAAGTGTGT 911
Db 3970 TCTCTACTGTACAGCAACTGTGGTTTCTCTCATCGAGCATCTCCAATTAAGTGTGT 3911
QY 912 CAGAGCATCTCTGCTTCTCAGACTGGCAGGCT 945
Db 3910 CAGAGGCTCTCTGCTTCTCTCACTGGCAGGCT 3877

RESULT 15
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-6

US-09-245-041-6

```
Query Match      3.1%; Score 132.2; DB 3; Length 226;
Best Local Similarity 78.6%; Pred. No. 2.4e-24;
Matches 158; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 2611 GATATGTCGCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGCT 2670
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 225  GATATGTCCTCCATTCACAAATAGTTTCTGCTGAGTGGATCCATCTGAGCCCACTGATGCT 166
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2671 GGATTCGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATC 2730
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 165  GGCTTCTGTGGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGCAACCTGCATC 106
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2731 AACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAACCCACAGTCTTAAGCAGTGCCCG 2790
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 105  AACCCCTCTCAATGGCAGCGTCTGTGAAGGCTGCTGAAGGACATGGGTGCATATAGTGCT 46
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2791 ACACCATGTGCTTGGAGCA 2811
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 45  CCAGGAGGAGCCACAGCA 25
```

Search completed: March 5, 2004, 12:20:34
Job time : 209.649 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2004, 12:05:12 ; Search time 15948 Seconds
(without alignments)
11659.244 Million cell updates/sec

Title: US-09-787-097-13

Perfect score: 4290

Sequence: 1 atgtgtccgcagcggcgc.....agctgggacctgcattctga 4290

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.cm.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3801	88.6	3819	9	AF106861	AF106861 Homo sapi
2	3774	88.0	8589	6	AR164816	AR164816 Sequence
3	3285	76.6	4072	6	AR164817	AR164817 Sequence
4	3045	71.0	3597	9	AF034957	AF034957 Homo sapi
5	1837	42.8	2825	6	AR164818	AR164818 Sequence
6	1697	39.6	2578	9	AK000356	AK000356 Homo sapi
7	1510	35.2	6370	6	AR164815	AR164815 Sequence
8	1361	31.7	5632	9	AB011120	AB011120 Homo sapi
9	543	12.7	543	6	AX778545	AX778545 Sequence
10	299	7.0	1576	9	H009384S01	AF218899 Homo sapi
11	299	7.0	51371	9	AL353193	AL353193 Human DNA
12	239	7.0	155869	2	AC023420	AC023420 Homo sapi
13	291	6.8	8827	6	AR164807	AR164807 Sequence
14	246	5.7	246	6	AR273223	AR273223 Sequence
15	246	5.7	246	6	AR276804	AR276804 Sequence
16	246	5.7	246	6	AR407079	AR407079 Sequence
17	246	5.7	246	6	AX368256	AX368256 Sequence
18	234	5.5	3257	9	H009384S15	AF218893 Homo sapi
19	234	5.5	104507	9	HSDJ741H3	AL132773 Human DNA
20	234	5.5	118540	2	AC013324	AC013324 Homo sapi
21	234	5.5	167357	2	AC055771	AC055771 Homo sapi
22	234	5.5	206647	9	AF002898	AF002898 Homo sapi
23	234	5.5	207433	6	AX326809	AX326809 Sequence
24	223	5.2	3589	9	H009384S11	AF218899 Homo sapi
25	223	5.2	155869	2	AC023420	AC023420 Homo sapi
26	221	5.2	1728	9	H009384S14	AF218892 Homo sapi
27	210	4.9	2018	9	H009384S08	AF218896 Homo sapi
28	193	4.5	63513	2	AC074061	AC074061 Homo sapi
29	192	4.5	3112	9	H009384S13	AF218901 Homo sapi
30	186	4.3	3038	9	H009384S09	AF218897 Homo sapi
31	173	4.0	1032	9	H009384S06	AF218894 Homo sapi
32	163	3.8	1232	9	H009384S05	AF218893 Homo sapi
33	161	3.8	1976	9	H009384S21	AF218909 Homo sapi
34	161	3.8	167357	2	AC055771	AC055771 Homo sapi
35	161	3.8	206329	2	AC017113	AC017113 Homo sapi
36	152	3.5	4493	4	AF531101	AF531101 Bos tauru
37	144	3.4	553	11	HS741H3S	AL110053 H. sapiens
38	141	3.3	1674	9	H009384S16	AF218904 Homo sapi
39	136	3.2	118540	2	AC013324	AC013324 Homo sapi
40	131	3.1	1152	9	H009384S04	AF218892 Homo sapi
41	131	3.1	73666	9	HSDJ581P3	AL109805 Human DNA
42	127	3.0	5790	9	H009384S27	AF218915 Homo sapi
43	127	3.0	206329	2	AC017113	AC017113 Homo sapi
44	119	2.8	1394	9	H009384S26	AF218914 Homo sapi
45	114	2.7	3058	9	H009384S03	AF218891 Homo sapi
46	113	2.6	237	4	AF273610	AF273610 Sus scrofa
47	110	2.6	1592	9	H009384S25	AF218913 Homo sapi
48	94	2.2	1675	9	H009384S17	AF218905 Homo sapi
49	92	2.1	1575	9	H009384S07	AF218895 Homo sapi
50	88	2.1	3711	9	H009384S02	AF218890 Homo sapi
51	86	2.0	1552	9	H009384S10	AF218898 Homo sapi
52	82	1.9	1556	9	H009384S24	AF218912 Homo sapi
53	79	1.8	1704	9	H009384S20	AF218908 Homo sapi
54	75	1.7	1267	9	H009384S12	AF218900 Homo sapi
55	65	1.5	2214	9	H009384S23	AF218911 Homo sapi
56	64	1.5	4284	10	AB062913	AB062913 Mesocricet
57	62	1.4	4313	10	AF119821	AF119821 Mus muscu
58	62	1.4	4500	10	AB038388	AB038388 Rattus no
59	62	1.4	5883	10	AB093245	AB093245 Mus muscu
60	62	1.4	6083	10	AF116897	AF116897 Mus muscu
61	62	1.4	8739	10	AB038387	AB038387 Rattus no
62	55	1.3	90050	6	AL64810	AL64810 Sequence
63	55	1.3	189283	10	AL833771	AL833771 Mouse DNA
64	53	1.2	966	9	H009384S18	AF218906 Homo sapi
65	53	1.2	1051	6	AR164814	AR164814 Sequence

[illegible][illegible]

DEFINITION		Sequence 14 from patent US 6274339.	
ACCESSION		AR164816	
VERSION		AR164816.1 GI:16238025	
KEYWORDS		Unknown.	
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		1 (bases 1 to 8589)	
AUTHORS		Moore, K. and Nagle, D. Lynn.	
TITLE		Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity	
JOURNAL		Patent: US 6274339-A 14 AUG 2001;	
FEATURES		Location/Qualifiers	
		1..8589	
		/organism="unknown"	
		/mol_type="unassigned DNA"	
ORIGIN		Query Match 88.0%; Score 3774; DB 6; Length 8589;	
		Best Local Similarity 99.9%; Pred. No. 0;	
		Matches 3974; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	313	TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCCAGTGGCTGTGCCCGCGGTGG	372
DB	310	TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCCAGTGGCTGTGCCCGCGGTGG	369
QY	373	GTGGCGAGCAATGCGAGCACTGCGGGGGCGCTTCAGACTAATCGATCTCTGGGTTT	432
DB	370	GTGGCGAGCAATGCGAGCACTGCGGGGGCGCTTCAGACTAATCGATCTCTGGGTTT	429
QY	433	GTGACAGATGACCTCGAAATTTATAATACAAACGAAAGTGCAGCTGCTCATTTGAAGGA	492
DB	430	GTGACAGATGACCTCGAAATTTATAATACAAACGAAAGTGCAGCTGCTCATTTGAAGGA	489
QY	493	CAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTCTACAGATGAGTGGAC	552
DB	490	CAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTCTACAGATGAGTGGAC	549
QY	553	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGC	612
DB	550	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGC	609
QY	613	CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTCACATCAGGT	672
DB	610	CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTCACATCAGGT	669
QY	673	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTTACCTGATTTTAAATTTACT	732
DB	670	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTTACCTGATTTTAAATTTACT	729
QY	733	TACAGTTTGTATGATGCTCCAAATTAATCTCAGCCCGAGGAGAGTGTAAAGATCAGTAAT	792
DB	730	TACAGTTTGTATGATGCTCCAAATTAATCTCAGCCCGAGGAGAGTGTAAAGATCAGTAAT	789
QY	793	AGCAGGAACTGTTGTAATGTAATGTTCTGAAACTGAAAGGTGAGCATGACATTT	852
DB	790	AGCAGGAACTGTTGTAATGTAATGTTCTGAAACTGAAAGGTGAGCATGACATTT	849
QY	853	CCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGCATCTGCAATTCAGATGATGTC	912
DB	850	CCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGCATCTGCAATTCAGATGATGTC	909
QY	913	AGAGATGCTCTGCTTCTCAGACTGGCAGGGTCTGATGTTTCTGATCTCTGATCCAGCT	972
DB	910	AGAGATGCTCTGCTTCTCAGACTGGCAGGGTCTGATGTTTCTGATCTCTGATCCAGCT	969
QY	973	AACCACTGATTTTGGACTCGAGAGGAATTTCTAACTTTAAAGCTCCCGAGCATCTCAT	1032
DB	970	AACCACTGATTTTGGACTCGAGAGGAATTTCTAACTTTAAAGCTCCCGAGCATCTCAT	1029
QY	1033	AAAGCTGTGCTCAATGGAAACATTTATGTTGGTGTGTTGAGGATATATGTTCAACCTCA	1092
DB	1030	AAAGCTGTGCTCAATGGAAACATTTATGTTGGTGTGTTGAGGATATATGTTCAACCTCA	1089

QY	1093	GATTATAACATGGTTCTAGCGTATGACCTTCTTAGGAGTGGCTTCCACTAAACCGT	1152
Db	1090	GATTATAACATGGTTCTAGCGTATGACCTTCTTAGGAGTGGCTTCCACTAAACCGT	1149
QY	1153	TCTGTGAACAAATGGTGTGTAGATATGTCATCTTTGGCATATACAAAGATAAAT	1212
Db	1150	TCTGTGAACAAATGGTGTGTAGATATGTCATCTTTGGCATATACAAAGATAAAT	1209
QY	1213	TACATGTATGAGGAAATTTGATCAACTGGGAATGTACCAATGATGTAGAGTTTTT	1272
Db	1210	TACATGTATGAGGAAATTTGATCAACTGGGAATGTACCAATGATGTAGAGTTTTT	1269
QY	1273	CACATTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1332
Db	1270	CACATTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1329
QY	1333	GTGGGCACTCTGCACATGTTTACATGGAAGTGGCGGCTGATGCTGGTCAATC	1392
Db	1330	GTGGGCACTCTGCACATGTTTACATGGAAGTGGCGGCTGATGCTGGTCAATC	1389
QY	1393	TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGATAAG	1452
Db	1390	TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGATAAG	1449
QY	1453	AACACATGAGTATATATACACACCCAGGGTGGCCCTTGTGCAAGGGGTTACGGCCATAGC	1512
Db	1450	AACACATGAGTATATATACACACCCAGGGTGGCCCTTGTGCAAGGGGTTACGGCCATAGC	1509
QY	1513	AGTGTATACACCATAGGACCGGCCCTATAGCTTCACTGGTGGCTTACAGGGCTTTCACT	1572
Db	1510	AGTGTATACACCATAGGACCGGCCCTATAGCTTCACTGGTGGCTTACAGGGCTTTCACT	1569
QY	1573	GCCAAATAGTACCGGCTTGAGATGATCTCTACCGATATGATGATGATGATGATGATG	1632
Db	1570	GCCAAATAGTACCGGCTTGAGATGATCTCTACCGATATGATGATGATGATGATGATG	1629
QY	1633	ACCATTTTAAAGACACCGCATTTTCCGTTACTGTCACACAGCTGTGATGATGATG	1692
Db	1630	ACCATTTTAAAGACACCGCATTTTCCGTTACTGTCACACAGCTGTGATGATGATG	1689
QY	1693	ACCATGTGTGTGTGGGGAACACACAAATGACACATCTATGAGCCATGCGCCAAA	1752
Db	1690	ACCATGTGTGTGTGGGGAACACACAAATGACACATCTATGAGCCATGCGCCAAA	1749
QY	1753	TGCTTCTTTCAGATTCAGGCTATGACATTCCTGTCACCGCTGGTGGCTTCCC	1812
Db	1750	TGCTTCTTTCAGATTCAGGCTATGACATTCCTGTCACCGCTGGTGGCTTCCC	1809
QY	1813	AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACACAGC	1872
Db	1810	AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACACAGC	1869
QY	1873	ACCATGTATGTGTGGTGTTCATATGATCTCTCTCTAGCGACATCTCTGGTATTCACC	1932
Db	1870	ACCATGTATGTGTGGTGTTCATATGATCTCTCTCTAGCGACATCTCTGGTATTCACC	1929
QY	1933	TCCGAACAGTGTGATGGCATCGAGTGAAGCGCTTGTATTAGCAGCAGCCTGTATT	1992
Db	1930	TCCGAACAGTGTGATGGCATCGAGTGAAGCGCTTGTATTAGCAGCAGCCTGTATT	1989
QY	1993	CGGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAA	2052
Db	1990	CGGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAA	2049
QY	2053	CAAGAGAAAGTTAAATCAGATGTTTTTCCAAAGACTCTTACCATGACAGATGT	2112
Db	2050	CAAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTACCATGACAGATGT	2109
QY	2113	GACCAGCACACAGATTTTACAGCTGTACAGCCAAACCAATGACCTGCCACTGGTCAAT	2172
Db	2110	GACCAGCACACAGATTTTACAGCTGTACAGCCAAACCAATGACCTGCCACTGGTCAAT	2169

QY	2173	GACCATTTGTGTCGCCAGGAACCAACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTAT	2232
Db	2170	GACCATTTGTGTCGCCAGGAACCAACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTAT	2229
QY	2233	GAGAAATGCCCCCAAGGATAACCCCATGTACTACTGTAAACAAAGAACAGCTGCAGGAGC	2292
Db	2230	GAGAAATGCCCCCAAGGATAACCCCATGTACTACTGTAAACAAAGAACAGCTGCAGGAGC	2289
QY	2293	TGTGCCCTTGGACCAAGAACTGCCAGTGGGAGCCCCCGAATCAGGAGTGCATTTGCCCTGCC	2352
Db	2290	TGTGCCCTTGGACCAAGAACTGCCAGTGGGAGCCCCCGAATCAGGAGTGCATTTGCCCTGCC	2349
QY	2353	GAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAACTCATGTTGAAATTTACTACT	2412
Db	2350	GAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAACTCATGTTGAAATTTACTACT	2409
QY	2413	GCCAAAGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAACAATGCCCTTTGGCT	2472
Db	2410	GCCAAAGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAACAATGCCCTTTGGCT	2469
QY	2473	TCTCTTACAAACCCAGAGAGTGAATTTCTTAAAGCAGCTGGGAATTAATGAGTCA	2532
Db	2470	TCTCTTACAAACCCAGAGAGTGAATTTCTTAAAGCAGCTGGGAATTAATGAGTCA	2529
QY	2533	TCTCAGACATGTCCAAGCTCACCTTAACCCCATGGTGGCCCTTCCGAAGATCAATGTG	2592
Db	2530	TCTCAGACATGTCCAAGCTCACCTTAACCCCATGGTGGCCCTTCCGAAGATCAATGTG	2589
QY	2593	TCCTACTGGTCTGGGAAGATATGTCCTTAAACAAATAGTTTACTACAGTGGATGCG	2652
Db	2590	TCCTACTGGTCTGGGAAGATATGTCCTTAAACAAATAGTTTACTACAGTGGATGCG	2649
QY	2653	TCTGAGCCAGTGTCTGGATTTCTGGAATTTTATCAGAACCCAGTACTCGGGGACTG	2712
Db	2650	TCTGAGCCAGTGTCTGGATTTCTGGAATTTTATCAGAACCCAGTACTCGGGGACTG	2709
QY	2713	AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTAAAGCCCTGCAAAACAC	2772
Db	2710	AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTAAAGCCCTGCAAAACAC	2769
QY	2773	AGTGTAAAGCAGTCCGACACACATGTCCTTGAAGCAGATGTGGAGATGACACAGC	2832
Db	2770	AGTGTAAAGCAGTCCGACACACATGTCCTTGAAGCAGATGTGGAGATGACACAGC	2829
QY	2833	GGCAGCTCTGAGTGTGATGTCAGCAACATGAAGCAGTGTGTGGACTCCAAATGCCCTAT	2892
Db	2830	GGCAGCTCTGAGTGTGATGTCAGCAACATGAAGCAGTGTGTGGACTCCAAATGCCCTAT	2889
QY	2893	GTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCTCC	2952
Db	2890	GTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCTCC	2949
QY	2953	GAAATTTCTCAGGCTACTGTACTGTAGTCTTGTGGAGCAACAGGCTGTGGCTGG	3012
Db	2950	GAAATTTCTCAGGCTACTGTACTGTAGTCTTGTGGAGCAACAGGCTGTGGCTGG	3009
QY	3013	TGTACTGTATCCAGCAATCTGGCAAGGAAATGATAGAGGGTTCCTATAAAGGACCA	3072
Db	3010	TGTACTGTATCCAGCAATCTGGCAAGGAAATGATAGAGGGTTCCTATAAAGGACCA	3069
QY	3073	GTGAAGATGCTTCCAGAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTC	3132
Db	3070	GTGAAGATGCTTCCAGAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTC	3129
QY	3133	AGCATGTGTCTAGAGGACAGAGATACAACTGGTCTTCTTCACTGTCCAGCTTGCCAA	3192
Db	3130	AGCATGTGTCTAGAGGACAGAGATACAACTGGTCTTCTTCACTGTCCAGCTTGCCAA	3189
QY	3193	TGCAACCGCCACAGTAAATGATCAATCAGAGATCTCTGTGAGAGTGTGAGAACTGACC	3252
Db	3190	TGCAACCGCCACAGTAAATGATCAATCAGAGATCTCTGTGAGAGTGTGAGAACTGACC	3249
QY	3253	ACAGGCAAGCATGTGGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGG	3312

Db 3250 ACAGGCAAGCACTGCGAGACTGCTCATATCTGGCTTTCTACGGTATCCCAATAGGAGG 3309
 QY 3313 AAATGTGAGCCATCAAGTGCATGCGGCGCTCTGTGCAACCAACCAAGCGGCAAG 3372
 Db 3310 AAATGTGAGCCATCAAGTGCATGCGGCGCTCTGTGCAACCAACCAAGCGGCAAG 3369
 QY 3373 TGTCTTCTGACCAACCAAGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAT 3432
 Db 3370 TGTCTTCTGACCAACCAAGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAT 3429
 QY 3433 CGATACCAAGAACCTCTCAGAGGAACATGTTATATATCTCTTCTTATGACTATCAG 3492
 Db 3430 CGATACCAAGAACCTCTCAGAGGAACATGTTATATATCTCTTCTTATGACTATCAG 3489
 QY 3493 TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCT 3552
 Db 3490 TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCT 3549
 QY 3553 ACTCTGACGAACAAACAGGGATTTGGAATGTTCAATCAATGCTCCAAAGAAATTTCAAC 3612
 Db 3550 ACTCTGACGAACAAACAGGGATTTGGAATGTTCAATCAATGCTCCAAAGAAATTTCAAC 3609
 QY 3613 CTCACATCACTGGGCTGCCAGTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCC 3672
 Db 3610 CTCACATCACTGGGCTGCCAGTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCC 3669
 QY 3673 GTTGTCTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAAGTTTGTAT 3732
 Db 3670 GTTGTCTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAAGTTTGTAT 3729
 QY 3733 TTTGCGAACCAACCAATATACATCTCTTCTTCTTATGTCAGTAATTCACCTGGCCATC 3792
 Db 3730 TTTGCGAACCAACCAATATACATCTCTTCTTCTTATGTCAGTAATTCACCTGGCCATC 3789
 QY 3793 AAAATTTCAGATTTGCTCTCTCAGCAGCAAAATTTATGACCTGGTACAGTTCTTTCG 3852
 Db 3790 AAAATTTCAGATTTGCTCTCTCAGCAGCAAAATTTATGACCTGGTACAGTTCTTTCG 3849
 QY 3853 ACTTCTTCAGTTGTTCT 3912
 Db 3850 ACTTCTTCAGTTGTTCT 3909
 QY 3913 CAAAGTTTGTGGGCTCCAGACGTAGAGAGCACTCTCTCAGAGATGCAACAGATGGCC 3972
 Db 3910 CAAAGTTTGTGGGCTCCAGACGTAGAGAGCACTCTCTCAGAGATGCAACAGATGGCC 3969
 QY 3973 AGCGTCCCTTTGCTCTGTAAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4032
 Db 3970 AGCGTCCCTTTGCTCTGTAAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4029
 QY 4033 ATTGGGGGAGTAAAGACTGTCTCCCAACCACTTCCCAACCACTTCCCAACCACTTCCCAAC 4092
 Db 4030 ATTGGGGGAGTAAAGACTGTCTCCCAACCACTTCCCAACCACTTCCCAACCACTTCCCAAC 4089
 QY 4093 AAAGCCGCTGTCT 4152
 Db 4090 AAAGCCGCTGTCT 4149
 QY 4153 CTTGGGAGTCAAGTCT 4212
 Db 4150 CTTGGGAGTCAAGTCT 4209
 QY 4213 ATAGTGTACAGAGAGTCAAGAGCCGTGAGAAACCGGAGCAGCAGCCCCCTGCACAG 4272
 Db 4210 ATAGTGTACAGAGAGTCAAGAGCCGTGAGAAACCGGAGCAGCAGCCCCCTGCACAG 4269
 QY 4273 CTTGGGAGTCAAGTCT 4290
 Db 4270 CTTGGGAGTCAAGTCT 4287

ARI64817
 LOCUS ARI64817 4072 bp DNA linear PAT 17-OCT-2001
 DEFINITION Sequence 16 from patent US 6274339.
 ACCESSION ARI64817
 VERSION ARI64817.1 GI:16238027
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4072)
 AUTHORS Moore, K. and Nagle, D. Lynn.
 TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
 JOURNAL Patent: US 6274339-A 16 14-AUG-2001;
 FEATURES Location/Qualifiers
 source 1..4072
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 76.6%; Score 3285; DB 6; Length 4072;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 313 TGTGTCAACGGCGGTGCGTGCACCCCTGGCACCGGCGGCTGCTGCCCGCGGCTGG 372
 Db 310 TGTGTCAACGGCGGTGCGTGCACCCCTGGCACCGGCGGCTGCTGCCCGCGGCTGG 369
 QY 373 GTGGCGAGCAATGCCAGCACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGTTT 432
 Db 370 GTGGCGAGCAATGCCAGCACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGTTT 429
 QY 433 GTGACAGATGACACCTGGAAATATAAATACAAACGAGTGCAGTGGCTCATTTGAAGGA 492
 Db 430 GTGACAGATGACACCTGGAAATATAAATACAAACGAGTGCAGTGGCTCATTTGAAGGA 489
 QY 493 CAGCCAAATAGAAATAGCACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGAC 552
 Db 490 CAGCCAAATAGAAATAGCACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGAC 549
 QY 553 CATTTATATGTTTATGATGGGACTCAATTTATGCAACGCTAGTGTCTGCAATTTAGTGGC 612
 Db 550 CATTTATATGTTTATGATGGGACTCAATTTATGCAACGCTAGTGTCTGCAATTTAGTGGC 609
 QY 613 CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTCAGAGTGTGTCACATCAGGT 672
 Db 610 CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTCAGAGTGTGTCACATCAGGT 669
 QY 673 TATGCTTTGCTGCAATTTTATGATGCTGCTTATAATTTGATGCTGATTTAATTTACT 732
 Db 670 TATGCTTTGCTGCAATTTTATGATGCTGCTTATAATTTGATGCTGATTTAATTTACT 729
 QY 733 TACAGTTTGTATGATGTGTCCTCAATTAATGCTCAGCCGAGAGAGTGTAGATCAGTAAT 792
 Db 730 TACAGTTTGTATGATGTGTCCTCAATTAATGCTCAGCCGAGAGAGTGTAGATCAGTAAT 789
 QY 793 AGCAGCGAACTGTTGAAATGTGAATGTTCTGAAATCTGAAAGTGAAGCATGTGACATT 852
 Db 790 AGCAGCGATACTGTTGAATGTGAATGTTCTGAAATCTGAAAGTGAAGCATGTGACATT 849
 QY 853 CCTCACTGTACAGCACTGTTGTTTCTCTCATCGAGCATCTGCAATTCAGGTGATGTC 912
 Db 850 CCTCACTGTACAGCACTGTTGTTTCTCTCATCGAGCATCTGCAATTCAGGTGATGTC 909
 QY 913 AGAGGATCTCTCTCTCTCTCAGACTGCGAGGCTCCTGATGTTTCTGTTTCTGTACAGCT 972
 Db 910 AGAGGATCTCTCTCTCTCTCAGACTGCGAGGCTCCTGATGTTTCTGTTTCTGTACAGCT 969
 QY 973 AACCACTATTTTGGACTCGAGAGAAATATTTCTAACTTAAAGTCTCCCGAGAGCATCTCAT 1032
 Db 970 AACCACTATTTTGGACTCGAGAGAAATATTTCTAACTTAAAGTCTCCCGAGAGCATCTCAT 1029
 QY 1033 AAAGCTGTGGTCAATGGAAACATTTATGTGGGTTTGTGGAGGATATATGTTCAACACTCA 1092

1030	DB	AAAGCTGTGTGCTCAATGTGAAACATTATATGTGGGTGTGTGGAGGATATATGTTTCAACACTCA	1089
1093	QY	GATTATAACATGTTCTTAGCGTATGACCTTGGCTTCTAGGGAGTGGCTTCCACTAAACCGT	1152
1090	DB	GATTATAACATGTTCTTAGCGTATGACCTTGGCTTCTAGGGAGTGGCTTCCACTAAACCGT	1149
1153	QY	TCTGTGAACAATGTGGTGTGTAGATATGGTCACTCTTTTGGCATATATACAAGGATAAAATT	1212
1150	DB	TCTGTGAACAATGTGGTGTGTAGATATGGTCACTCTTTTGGCATATATACAAGGATAAAATT	1209
1213	QY	TACATGTATCGAGGAAATAATGTATCCAACTGGGAACTGTGACCAATGAGTTGACAGTTTTTT	1272
1210	DB	TACATGTATCGAGGAAATAATGTATCCAACTGGGAACTGTGACCAATGAGTTGACAGTTTTTT	1269
1273	QY	CACATTCATAATGAGTCATGGGTGTGTGTGACCCCTAAGGCAAGGACGAGTATGCAGTG	1332
1270	DB	CACATTCATAATGAGTCATGGGTGTGTGTGACCCCTAAGGCAAGGACGAGTATGCAGTG	1329
1333	QY	GTTGGGCACCTCGCACACATGTGTACCTGAAGNATGGCCGAGTGTCTATGCTGGTCATC	1392
1330	DB	GTTGGGCACCTCGCACACATGTGTACCTGAAGNATGGCCGAGTGTCTATGCTGGTCATC	1389
1393	QY	TTTGGTCACCTGCCCTCTCTATGATATATAAGCAATGTGACGAATAATGATTTGGGATAAG	1452
1390	DB	TTTGGTCACCTGCCCTCTCTATGATATATAAGCAATGTGACGAATAATGATTTGGGATAAG	1449
1453	QY	AACACATGGAGTATATTACACACCCAGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGC	1512
1450	DB	AACACATGGAGTATATTACACACCCAGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGC	1509
1513	QY	AGTGTTTACCAACCATAGGACACAGGGCCCTATACGTTTCATGGTGGCTACAAAGGCTTTCAGT	1572
1510	DB	AGTGTTTACCAACCATAGGACACAGGGCCCTATACGTTTCATGGTGGCTACAAAGGCTTTCAGT	1569
1573	QY	GCCAAATAGTACCGGCTTGAGATGATCTCTACCGATATGATGFGNATACCCAGATGTGG	1632
1570	DB	GCCAAATAGTACCGGCTTGAGATGATCTCTACCGATATGATGFGNATACCCAGATGTGG	1629
1633	QY	ACCAATTCCTTAAGGACACCGCATTTTCCGTTTACTTTGCACACAGCTGTGATGTAGTGGA	1692
1630	DB	ACCAATTCCTTAAGGACACCGCATTTTCCGTTTACTTTGCACACAGCTGTGATGTAGTGGA	1689
1693	QY	ACCATGCTGTGTGTGGGGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAA	1752
1690	DB	ACCATGCTGTGTGTGGGGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAA	1749
1753	QY	TGCTTCTCTTCAGATTTCAAGCCCTATGACATTCGCTGTGACCGCTGGTCAGTGTCTCC	1812
1750	DB	TGCTTCTCTTCAGATTTCAAGCCCTATGACATTCGCTGTGACCGCTGGTCAGTGTCTCC	1809
1813	QY	AGACCTGATCTCCACCATGATGTCAAAGATTTGGCCATTCAGCAGTCTTTACACAAACGC	1872
1810	DB	AGACCTGATCTCCACCATGATGTCAAAGATTTGGCCATTCAGCAGTCTTTACACAAACGC	1869
1873	QY	ACCATGATATGTTTCCGTTGGTTTTCAATAGTCTCTCTCTCAGCGACATCTCTGGTATTCAC	1932
1870	DB	ACCATGATATGTTTCCGTTGGTTTTCAATAGTCTCTCTCTCAGCGACATCTCTGGTATTCAC	1929
1933	QY	TCGGNACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTTCAGCAGCAGGACCTGGTATT	1992
1930	DB	TCGGNACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTTCAGCAGCAGGACCTGGTATT	1989
1993	QY	CGGTGTGTGCGAAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAA	2052
1990	DB	CGGTGTGTGCGAAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAA	2049
2053	QY	CAAGNAGAAAGTTAAATCAGATGTTTTTTTCCAAAAGAACTCTTTGACCATGACAGATGT	2112
2050	DB	CAAGNAGAAAGTTAAATCAGATGTTTTTTTCCAAAAGAACTCTTTGACCATGACAGATGT	2109
2113	QY	GACCAGACACAGATTTGTATACAGCTGTATACGCCAACCAATGATGTGCCACTGGTGCAT	2172

Db	2110	GACCAGCACACAGATTGTTACAGCTGCACAGCCACACCAATGACTGCCACTGGTGC	2116
Qy	2173	GACCAATTGTTCCCCAGGAACCA	2232
Db	2170	GACCAATTGTTCCCCAGGAACCA	2229
Qy	2233	GAGNAATTGCCCCCAAGNATACCCCATCTACTACTGTAACAAAGAACCAAGCTGCAGGAGC	2292
Db	2230	GAGAAATTGCCCCCAAGGATAACCCATATGTACTCTGTAAAGAAAGACCAAGCTGCAGGAGC	2289
Qy	2293	TGTGCCCTGGACAGCAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2352
Db	2290	TGTGCCCTGGACAGCAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2349
Qy	2353	GAAATATCTGTGGCATTGGCTGGCAATTTGGTTGGAAACTCATGTTTGAATAATTACTACT	2412
Db	2350	GAAATATCTGTGGCATTGGCTGGCAATTTGGTTGGAAACTCATGTTTGAATAATTACTACT	2409
Qy	2413	GCCAAAGAGAAATTATGACAAATGCTAAATTTGTTGTAGGAACACAAATGCCCTTTTGGCT	2472
Db	2410	GCCAAAGAGAAATTATGACAAATGCTAAATTTGTTGTAGGAACACCAATGCCCTTTTGGCT	2469
Qy	2473	TCTCTTACAACCCAGAGAAAGTATGAAATTTGTTCTTAAGCAGCTGCAGAAATAATGCAGTCA	2532
Db	2470	TCTCTTACAACCCAGAGAAAGTATGAAATTTGTTCTTAAGCAGCTGCAGAAATAATGCAGTCA	2529
Qy	2533	TCTCAGAGCATGTCCTTAACCCCATGGCTCGGCCCTTCGGAAGATCAATGCTG	2592
Db	2530	TCTCAGAGCATGTCCTTAACCCCATGGCTCGGCCCTTCGGAAGATCAATGCTG	2589
Qy	2593	TCCTACTGGTCTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATGCCG	2652
Db	2590	TCCTACTGGTCTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATGCCG	2649
Qy	2653	TCTCAGCCAGTGTGCTGGATTCGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTG	2712
Db	2650	TCTCAGCCAGTGTGCTGGATTCGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTG	2709
Qy	2713	AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGCCCTGC	2772
Db	2710	AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGCCCTGC	2769
Qy	2773	AGTGCTAAGCAGTCCGGACACCAATGTCCTTGAGGACAGCATGTGGAGATTCAC	2832
Db	2770	AGTGCTAAGCAGTCCGGACACCAATGTCCTTGAGGACAGCATGTGGAGATTCAC	2829
Qy	2833	GGCAGCTCTGAGTGCATGTGGTGACGAAATGAAAGCAGTGTGAGACTCCAAATGCCTAT	2892
Db	2830	GGCAGCTCTGAGTGCATGTGGTGACGAAATGAAAGCAGTGTGAGACTCCAAATGCCTAT	2889
Qy	2893	GTGGCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGGACCTGCGCCCTT	2952
Db	2890	GTGGCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGGACCTGCGCCCTT	2949
Qy	2953	GAAATTTGTTCCAGCTACTGTGACCTGTAGTCAATTTGAGGAAACAGGCTGTGGCTGG	3012
Db	2950	GAAATTTGTTCCAGCTACTGTGACCTGTAGTCAATTTGAGGAAACAGGCTGTGGCTGG	3009
Qy	3013	TGTACTGATCCAGCAATCTGCGAAAGGGAATTCATAGAGGTTTCTATPAAAGGACCA	3072
Db	3010	TGTACTGATCCAGCAATCTGCGAAAGGGAATTCATAGAGGTTTCTATPAAAGGACCA	3069
Qy	3073	GTGAGATGCTTTCGACAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC	3132
Db	3070	GTGAGATGCTTTCGACAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC	3129
Qy	3133	AGCATGTGCTAGAGGACAGCAGATACAACTGGTCTTTTCAATTCAGTCTCCAGCTGGCAA	3192
Db	3130	AGCATGTGCTAGAGGACAGCAGATACAACTGGTCTTTTCAATTCAGTCTCCAGCTGGCAA	3189
Qy	3193	TGCAACGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACC	3252
Db	3190	TGCAACGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACC	3249

QY 3253 ACAGCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGATCCACCAATGGAGG 3312
Db 3250 ACAGCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGATCCACCAATGGAGG 3309
QY 3313 AAATGTACGACCAATGCAATGGGACGCGTCTCTGTGCAACCAACAGGCGCAAG 3372
Db 3310 AAATGTACGACCAATGCAATGGGACGCGTCTCTGTGCAACCAACAGGCGCAAG 3369
QY 3373 TGCCTTCTGACCAACAGGCGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAT 3432
Db 3370 TGCCTTCTGACCAACAGGCGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAT 3429
QY 3433 CGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTCTTATTGACTATCAG 3492
Db 3430 CGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTCTTATTGACTATCAG 3489
QY 3493 TTACCTTTAGTCTATCCAGGAAGATGATGCTTATTACAGTATCAATTTTGGCT 3552
Db 3490 TTACCTTTAGTCTATCCAGGAAGATGATGCTTATTACAGTATCAATTTTGGCT 3549
QY 3553 ACTCCTGACCAACCAACAGGATTTGCAATGTTTCAATCAATGCTCCAGAAATTTCAAC 3612
Db 3550 ACTCCTGACCAACCAACAGGATTTGCAATGTTTCAATCAATGCTCCAGAAATTTCAAC 3609
QY 3613 CTCACATACCTGGCTGCGAGTTTCTACGTGGAACCCAGGCTGAGAGAGATGCT 3672
Db 3610 CTCACATACCTGGCTGCGAGTTTCTACGTGGAACCCAGGCTGAGAGAGATGCT 3669
QY 3673 GTTCTTTCAAAACCAACATTAAGGATCAAAAGATGTTCTCTAATGAGAGTTTGTAT 3732
Db 3670 GTTCTTTCAAAACCAACATTAAGGATCAAAAGATGTTCTCTAATGAGAGTTTGTAT 3729
QY 3733 TTTGCGAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTCACTGGCCCATC 3792
Db 3730 TTTGCGAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTCACTGGCCCATC 3789
QY 3793 AAAATTGAG 3801
Db 3790 AAAATTGAG 3798

RESULT 4
AF034957
LOCUS AF034957.1 GI:3676346
DEFINITION Homo sapiens secreted T cell activation protein Attractin
(attractin) mRNA, complete cds.
ACCESSION AF034957
VERSION AF034957.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3597)
Duke-Cohan, J.S., Gu, J., McLaughlin, D.F., Xu, Y., Freeman, G.J. and
Schlossman, S.F.
TITLE Attractin (pppT-L), a member of the CUB family of cell adhesion and
guidance proteins, is secreted by activated human T lymphocytes and
modulates immune cell interactions
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
MEDLINE 98409658
PUBMED 9736737
REFERENCE 2 (bases 1 to 3597)
Duke-Cohan, J.S., Gu, J., McLaughlin, D.F., Freeman, G.J. and
Schlossman, S.F.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS,
Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
Location/Qualifiers
1..3597
/organism="Homo sapiens"

Query Match 71.0%; Score 3045; DB 9; Length 3597;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 312 CTGTGTCAACGGCGCTGCTGCAACCCCTGGCACCGGCGAGTGCCTGTCGCCCGCGGCTG 371
Db 90 CTGTGTCAACGGCGCTGCTGCAACCCCTGGCACCGGCGAGTGCCTGTCGCCCGCGGCTG 149
QY 372 GGTGGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAATCTGATCTTCTGGTT 431
Db 150 GGTGGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAATCTGATCTTCTGGTT 209
QY 432 TGTGACAGTGGACCTGGAATTAATAACAAACGAAGTGACGCTGGCTCATTTGAAGG 491
Db 210 TGTGACAGTGGACCTGGAATTAATAACAAACGAAGTGACGCTGGCTCATTTGAAGG 269
QY 492 ACAGCAAAATAGATAATGAGACTTGGTTTCAATCAATTTTCTACAGAGTGTAGTGGGA 551
Db 270 ACAGCAAAATAGATAATGAGACTTGGTTTCAATCAATTTTCTACAGAGTGTAGTGGGA 329
QY 552 CCATTTATCTTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGCTCATTTAGTGG 611
Db 330 CCATTTATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGCTCATTTAGTGG 389
QY 612 CCTCATTTCTCTCAGAGAGATGCAATGAGACTGTCTCCCTGAGGTTGTGCCACATCAGG 671
Db 390 CCTCATTTCTCTCAGAGAGATGCAATGAGACTGTCTCCCTGAGGTTGTGCCACATCAGG 449
QY 672 TTATGCTTCTGCTGATTTTCTGATGCTGCTTAAATTTGACTGATTTATATTAT 731
Db 450 TTATGCTTCTGCTGATTTTCTGATGCTGCTTAAATTTGACTGATTTATATTAT 509
QY 732 TTACAGTTTTGATATGTCTCAAAATAACTGCTCAGGCGGAGGAGTGAAGTACAGTAA 791
Db 510 TTACAGTTTTGATATGTCTCAAAATAACTGCTCAGGCGGAGGAGTGAAGTACAGTAA 569
QY 792 TAGCAGGAAACTGTTGAATGTGAATGTTCTGAAAACCTGGAAGGTGAAGCATGTGACAT 851

ORIGIN
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3597
/gene="attractin"
1..3597
/gene="attractin"
/note="secreted T cell activation protein; Method:
conceptual translation with partial peptide sequencing;
pppT-L"
/codon_start=1
/product="Attractin"
/protein_id="AAC61902.1"
/db_xref="GI:3676347"

QY	3011	GGTGTACTGATCCAGCAATACTGCGAAAGGAAATGATAGAGGTTCTTATTAAGGAC	3070
Db	2789	GGTGTACTGATCCAGCAATACTGCGAAAGGAAATGATAGAGGTTCTTATTAAGGAC	2848
QY	3071	CAGTCAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTATCCACGCCCTGCTCAATT	3130
Db	2849	CAGTCAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTATCCACGCCCTGCTCAATT	2908
QY	3131	CCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGGC	3190
Db	2909	CCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGGC	2968
QY	3191	AATGCAAGGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGAACCTGA	3250
Db	2969	AATGCAAGGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGAACCTGA	3028
QY	3251	CCAGGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCCAACCAATGGAG	3310
Db	3029	CCAGGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCCAACCAATGGAG	3088
QY	3311	GGAAATGTCAGCCATGCAAGTGCATGAGGCGCATCTCTGTGCAACACCAACACGGGCA	3370
Db	3089	GGAAATGTCAGCCATGCAAGTGCATGAGGCGCATCTCTGTGCAACACCAACACGGGCA	3148
QY	3371	AGTGTCTCTGCACCAACCAAGGGCGTCAAGGGGACGAGTGCCAGCTATGTGAGGTAGAAA	3430
Db	3149	AGTGTCTCTGCACCAACCAAGGGCGTCAAGGGGACGAGTGCCAGCTATGTGAGGTAGAAA	3208
QY	3431	ATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTAATCTCTTTTATGACTATC	3490
Db	3209	ATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTAATCTCTTTTATGACTATC	3268
QY	3491	AGTTCAACCTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATCAATTTTGTGG	3550
Db	3269	AGTTCAACCTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATCAATTTTGTGG	3328
QY	3551	CTACTCTCTGAGCAACAAACAGGGATTTGACATGTTTCATCAATGCCCTCCAGAAATTTCA	3610
Db	3329	CTACTCTCTGAGCAACAAACAGGGATTTGACATGTTTCATCAATGCCCTCCAGAAATTTCA	3388
QY	3611	ACCTCAACATCACTGGGCTCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGATGC	3670
Db	3389	ACCTCAACATCACTGGGCTCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGATGC	3448
QY	3671	CTGTGTTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTG	3730
Db	3449	CTGTGTTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTG	3508
QY	3731	ATTTTGGCAACCCCAAAATATCACTTCTTTTGTATGTCAGTAATTTCACTGGCCCA	3790
Db	3509	ATTTTGGCAACCCCAAAATATCACTTCTTTTGTATGTCAGTAATTTCACTGGCCCA	3568
QY	3791	TCAAAATTCAG 3801	
Db	3569	TCAAAATTCAG 3579	
RESULT 5			
LOCUS	AR164818	2625 bp	DNA
DEFINITION	Sequence 18 from patent US 6274339.	linear	PAT 17-OCT-2001
ACCESSION	AR164818		
VERSION	AR164818.1		GI:16238029
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2625)		
AUTHORS	Moore,K. and Nagle,D.Lynn.		
TITLE	Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity		
JOURNAL	Patent: US 6274339-A 18 14-AUG-2001;		

FEATURES	Location/Qualifiers	
source	1..2625	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	42.8%; Score 1837; DB 6; Length 2625;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 2037; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
QY	313	TGTGTCAACGGCGGTGCGTGCACACCTGGCACCGGCCAGTGTGCGCCCGCGGTGG 372
Db	310	TGTGTCAACGGCGGTGCGTGCACACCTGGCACCGGCCAGTGTGCGCCCGCGGTGG 369
QY	373	GTGGGCGAGCAATGCGAGCATGTGGGGGGCGCTTCAGACTAACTGGATCTCTGGGTTT 432
Db	370	GTGGGCGAGCAATGCGAGCATGTGGGGGGCGCTTCAGACTAACTGGATCTCTGGGTTT 429
QY	433	GTGACAGATGGACCTGGAAATTTATAATAACAAAGAGTGCACGTGGCTCATTTGAAGGA 492
Db	430	GTGACAGATGGACCTGGAAATTTATAATAACAAAGAGTGCACGTGGCTCATTTGAAGGA 489
QY	493	CAGCCAAATAGAAATATGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTAGTTGGAC 552
Db	490	CAGCCAAATAGAAATATGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTAGTTGGAC 549
QY	553	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGTGTCATTTAGTGGC 612
Db	550	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGTGTCATTTAGTGGC 609
QY	613	CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 672
Db	610	CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 669
QY	673	TATGCTTGTGTCATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAATTAAT 732
Db	670	TATGCTTGTGTCATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAATTAAT 729
QY	733	TACAGTTTGTATGATGCTCCAAATAAATGCTCAGGCCGAGGAGATGTAAGATCAGTAAT 792
Db	730	TACAGTTTGTATGATGCTCCAAATAAATGCTCAGGCCGAGGAGATGTAAGATCAGTAAT 789
QY	793	AGCAGCAAACTGTTGAATGTAATGTTCTGAAATACTGAAAGTGAAGCATGTGACATTT 852
Db	790	AGCAGCAAACTGTTGAATGTAATGTTCTGAAATACTGAAAGTGAAGCATGTGACATTT 849
QY	853	CCTCACTGTACAGACAATGTTGTTTCTCATCGAGGCATCTGCAATTCAGATGATGTC 912
Db	850	CCTCACTGTACAGACAATGTTGTTTCTCATCGAGGCATCTGCAATTCAGATGATGTC 909
QY	913	AGAGATGCTCCTGCTTCTCAGACTGGCAGGGTCTGATGTTTCACTGTTTCTGTTACCACT 972
Db	910	AGAGATGCTCCTGCTTCTCAGACTGGCAGGGTCTGATGTTTCACTGTTTCTGTTACCACT 969
QY	973	AACCACTCATTTTGGACTCGAGAGAAATTTCTAATTAAGCTCCCGAGAGCATCTCAT 1032
Db	970	AACCACTCATTTTGGACTCGAGAGAAATTTCTAATTAAGCTCCCGAGAGCATCTCAT 1029
QY	1033	AAAGCTGTGTCATGGAACCAATTAATGTTGGGTTGTTGGAGGATATATGTTCAACCACTCA 1092
Db	1030	AAAGCTGTGTCATGGAACCAATTAATGTTGGGTTGTTGGAGGATATATGTTCAACCACTCA 1089
QY	1093	GATTATACATGTTTCTAGCGTATGACCTTCTTAGGGAGTGGCTTCCACTAAACCGT 1152
Db	1090	GATTATACATGTTTCTAGCGTATGACCTTCTTAGGGAGTGGCTTCCACTAAACCGT 1149
QY	1153	TCTGTGAACCAATGTTGTTAGATATGCTCATTTCTTTGGCAATTATACAAGGATAAAAT 1212
Db	1150	TCTGTGAACCAATGTTGTTAGATATGCTCATTTCTTTGGCAATTATACAAGGATAAAAT 1209
QY	1213	TACATGTTAGGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT 1272
Db	1210	TACATGTTAGGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT 1269

```
QY 1273 CACATTCATATGAGTCATGGGTGTTGTTGACCCCTAAGSCRAAGCAGCATATGCAAGTG 1332
Db 1270 CACATTCATATGAGTCATGGGTGTTGTTGACCCCTAAGSCRAAGCAGCATATGCAAGTG 1329
QY 1333 GTTGGGCACCTCTGCACACATTTGTACACTGAAGATGGCCGAGTGTCTATGCTGGTGCATC 1392
Db 1330 GTTGGGCACCTCTGCACACATTTGTACACTGAAGATGGCCGAGTGTCTATGCTGGTGCATC 1389
QY 1393 TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGCAATATGATTTGGATAAG 1452
Db 1390 TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGCAATATGATTTGGATAAG 1449
QY 1453 AACACATGAGTATTTACACACCCAGGGTGCCTTTGTGCAAGGGGGTTACGGCCCATAGC 1512
Db 1450 AACACATGAGTATTTACACACCCAGGGTGCCTTTGTGCAAGGGGGTTACGGCCCATAGC 1509
QY 1513 AGTGTTTACACCATAGGACACAGGCCCTTATAGTTTCATGGTGGCTACAAAGCTTTTCAGT 1572
Db 1510 AGTGTTTACACCATAGGACACAGGCCCTTATAGTTTCATGGTGGCTACAAAGCTTTTCAGT 1569
QY 1573 GCCAATAAGTACCGGCTTGACAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGG 1632
Db 1570 GCCAATAAGTACCGGCTTGACAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGG 1629
QY 1633 ACCATTCTTAAGACACCCGATTTTCCGTTTACTTGCACACAGCTGTGATGATGTGGA 1692
Db 1630 ACCATTCTTAAGACACCCGATTTTCCGTTTACTTGCACACAGCTGTGATGATGTGGA 1689
QY 1693 ACCATGCTGTGTTTGGGGGAAACACACAAATGACATCTATGAGCCATGCGGCCAA 1752
Db 1690 ACCATGCTGTGTTTGGGGGAAACACACAAATGACATCTATGAGCCATGCGGCCAA 1749
QY 1753 TGCCTTCCTTCAGATTTTCATGGCCCTATGACATTTGCTGTGACCGCTGGTCAGTCTTCCC 1812
Db 1750 TGCCTTCCTTCAGATTTTCATGGCCCTATGACATTTGCTGTGACCGCTGGTCAGTCTTCCC 1809
QY 1813 AGACCTGATCTCCACCATGATGTCACAGATTTGGCCCATTCAGCAGTCTTACACACAGC 1872
Db 1810 AGACCTGATCTCCACCATGATGTCACAGATTTGGCCCATTCAGCAGTCTTACACACAGC 1869
QY 1873 ACCATGATGTGTTTCGGTGGTTTCAATAGTCTCTCTCCTCAGCACATCTCTGATTAACCC 1932
Db 1870 ACCATGATGTGTTTCGGTGGTTTCAATAGTCTCTCTCCTCAGCACATCTCTGATTAACCC 1929
QY 1933 TCGGAACAGTGTGATGCGCATTCGGAGTGAAGCCGCTTTTATAGCAGCAGGACCTGGTATT 1992
Db 1930 TCGGAACAGTGTGATGCGCATTCGGAGTGAAGCCGCTTTTATAGCAGCAGGACCTGGTATT 1989
QY 1993 CGGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGGCGTGGCACTGTATGAA 2052
Db 1990 CGGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGGCGTGGCACTGTATGAA 2049
QY 2053 CAAGAAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGT 2112
Db 2050 CAAGAAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGT 2109
QY 2113 GACCAAGACACAGATTTGATCAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCAT 2172
Db 2110 GACCAAGACACAGATTTGATCAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCAT 2169
QY 2173 GACCAATGTGTCCCAGGAACACAGCTGTCTCAGAAGGCCAGATCTCCATTTTAGGTAT 2232
Db 2170 GACCAATGTGTCCCAGGAACACAGCTGTCTCAGAAGGCCAGATCTCCATTTTAGGTAT 2229
QY 2233 GAGAATTTGCCCAAGGATAACCCCATGTACTGTATTAAGAAAGACAGCTGCAGGAGC 2292
Db 2230 GAGAATTTGCCCAAGGATAACCCCATGTACTGTATTAAGAAAGACAGCTGCAGGAGC 2289
QY 2293 TGTGCCCTGGACCAAGAACTGCAGTGGAGCCCGGGAATCAGGATGCAATTTGCCCTGCC 2352
Db 2290 TGTGCCCTGGACCAAGAACTGCAGTGGAGCCCGGGAATCAGGATGCAATTTGCCCTGCC 2349
```

```
QY 2353 G 2353
Db 2350 G 2350

RESULT 6
AK000356 2578 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA FLJ20349 fis, clone HRP13957, highly similar to
DEFINITION AK000356.1 GI:7020387
ACCESSION AK000356
VERSION cligo capping; fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,
Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2. (bases 1 to 2578)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
Minato-ku, Tokyo 108-8639, Japan (E-mail: f.cdn@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction; 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
1..2578
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP13957"
/cell_line="HEPG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="cloning vector pME18SFL3"
misc_feature
1..2578
/note="highly similar to AB011120 Homo sapiens mRNA for
KIAA0548 protein"
ORIGIN
Query Match 39.6%; Score 1697; DB 9; Length 2578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2594 CCTACTGTGCTCGGAAGATATGTCCTGGAATTTATCAGAACTCTCGGGGACTGA 2713
Db 1 CCTACTGTGCTCGGAAGATATGTCCTGGAATTTATCAGAACTCTCGGGGACTGA 2653
QY 2654 CTGAGCCCACTGATGCTGGATTTCTGGAATTTATCAGAACTCTCGGGGACTGA 120
Db 61 CTGAGCCCACTGATGCTGGATTTCTGGAATTTATCAGAACTCTCGGGGACTGA 120
QY 2714 AGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAACCA 2773
Db 121 AGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAACCA 180
QY 2774 GTGCTAAGCAGTGGCGGACACCATGTGCTTGCAGACAGCATGTGAGATTGCACAGCG 2833
Db 181 GTGCTAAGCAGTGGCGGACACCATGTGCTTGCAGACAGCATGTGAGATTGCACAGCG 240
```

```
QY 2834 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCCCTATG 2893
Db 241 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCCCTATG 300
QY 2894 TGGGCTCCCTTCCCTTTTGGCCAGTGTATGGATGGTATAGATGAGCACCTGCCCTCG 2953
Db 301 TGGGCTCCCTTCCCTTTTGGCCAGTGTATGGATGGTATAGATGAGCACCTGCCCTCG 360
QY 2954 AAAATTGTTTCAGGCTACTGTACTCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGGT 3013
Db 361 AAAATTGTTTCAGGCTACTGTACTCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGGT 420
QY 3014 GTACTGATCCAGCAATCTCTGGCAAGGAAATGCAATAGAGGTTTCTTAAGAGGACAG 3073
Db 421 GTACTGATCCAGCAATCTCTGGCAAGGAAATGCAATAGAGGTTTCTTAAGAGGACAG 480
QY 3074 TGAAGATGCTTCCAGGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTTCCA 3133
Db 481 TGAAGATGCTTCCAGGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTTCCA 540
QY 3134 GCATGTGTCTAGAGCAAGAGATACAACTGGTCTTTCAATCTCACTGCCAGTTGCCAAT 3193
Db 541 GCATGTGTCTAGAGCAAGAGATACAACTGGTCTTTCAATCTCACTGCCAGTTGCCAAT 600
QY 3194 GCAAGGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA 3253
Db 601 GCAAGGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA 660
QY 3254 CAGGCAAGCACTGGAGACCTGTCATATCTGGCTTTCTAGCTGTATGCCCAATGGAGGGA 3313
Db 661 CAGGCAAGCACTGGAGACCTGTCATATCTGGCTTTCTAGCTGTATGCCCAATGGAGGGA 720
QY 3314 AATGTACGATGCAAGTGAATGGGACCGTCTCTGTGCAACACCAACAGGGGCAAGT 3373
Db 721 AATGTACGATGCAAGTGAATGGGACCGTCTCTGTGCAACACCAACAGGGGCAAGT 780
QY 3374 GTTCTTGACCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGAGTGAAGAAATC 3433
Db 781 GTTCTTGACCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGAGTGAAGAAATC 840
QY 3434 GATACCAAGGAAACCTCTCAGAGGAACATGTATTATATCTCTTTATTGACTATCAGT 3493
Db 841 GATACCAAGGAAACCTCTCAGAGGAACATGTATTATATCTCTTTATTGACTATCAGT 900
QY 3494 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTACAGCTATCAATTTTGTGGCTA 3553
Db 901 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTACAGCTATCAATTTTGTGGCTA 960
QY 3554 CTCCTGAGGAACAAACAGGGAATTTGACATGTTTCATCAATGCTCCCAAGAAATTTCAACC 3613
Db 961 CTCCTGAGGAACAAACAGGGAATTTGACATGTTTCATCAATGCTCCCAAGAAATTTCAACC 1020
QY 3614 TCACATCAGCTGGCTGCCAGTTTCTCAGTGGAAACCCAGGCTGGAGAGAGATGCTG 3673
Db 1021 TCACATCAGCTGGCTGCCAGTTTCTCAGTGGAAACCCAGGCTGGAGAGAGATGCTG 1080
QY 3674 TTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATT 3733
Db 1081 TTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATT 1140
QY 3734 TTGCAACACCCAAATATCACTTTCTTTGTTATGTAGTAAATTTTCACTGCCCATCA 3793
Db 1141 TTGCAACACCCAAATATCACTTTCTTTGTTATGTAGTAAATTTTCACTGCCCATCA 1200
QY 3794 AAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCTGTA 3853
Db 1201 AAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCTGTA 1260
QY 3854 CTTTCTTCAGTTGTTTCTCTCTTTGCTCTGTTGGTCTGTTGGTTTGAAGATCAAAAC 3913
Db 1261 CTTTCTTCAGTTGTTTCTCTCTTTGCTCTGTTGGTCTGTTGGTTTGAAGATCAAAAC 1320
```

```
QY 3914 AAAGTTGTTGGCTCCAGACCTGAGAGCAAACTTTCTTCGAGAGATGCAACAGATGGCCA 3973
Db 1321 AAAGTTGTTGGCTCCAGACCTGAGAGCAAACTTTCTTCGAGAGATGCAACAGATGGCCA 1380
QY 3974 GCGCTCCCTTTCCCTCTGTAATGTCGCTTGGAAACAGATGAGAGCTCTCTGATCTTA 4033
Db 1381 GCGCTCCCTTTCCCTCTGTAATGTCGCTTGGAAACAGATGAGAGCTCTCTGATCTTA 1440
QY 4034 TTGGGGGAGTATAAAGACTGTTCCTCCAAACCCATTGCACTGGAGCCGTGTTTGGCAACA 4093
Db 1441 TTGGGGGAGTATAAAGACTGTTCCTCCAAACCCATTGCACTGGAGCCGTGTTTGGCAACA 1500
QY 4094 AAGCGCTGCTCTCTCTGTTGTGGGCTCCCTCGAGGCTGGTGGCATCCCTCTCTC 4153
Db 1501 AAGCGCTGCTCTCTCTGTTGTGGGCTCCCTCGAGGCTGGTGGCATCCCTCTCTC 1560
QY 4154 CTGGGCACTCAGCTCTGCTGTGGCCAGCGCTGGTGGACATTTCTCAGCAGATGCCGA 4213
Db 1561 CTGGGCACTCAGCTCTGCTGTGGCCAGCGCTGGTGGACATTTCTCAGCAGATGCCGA 1620
QY 4214 TAGTGTACAGGAGAGTCAAGAGCCGTGAGAACCCGAGCAGCCCTGCAAGC 4273
Db 1621 TAGTGTACAGGAGAGTCAAGAGCCGTGAGAACCCGAGCAGCCCTGCAAGC 1680
QY 4274 CTGGGACCTGCACTGA 4290
Db 1681 CTGGGACCTGCACTGA 1697

RESULT 7
AR164815 LOCUS AR164815 6370 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815
VERSION AR164815.1 GI:16238024
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6370)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 12 14-AUG-2001;
FEATURES Location/Qualifiers
source 1..6370
/organism="unknown"
/moi_type="unassigned DNA"

ORIGIN
Query Match 35.2%; Score 1510; DB 6; Length 6370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2781 CGAGTCCCGGACACCATGTGCTTGGAGCAGCAGCATGTGGAGATTGCACAGCGGCGATC 2840
Db 548 CGAGTCCCGGACACCATGTGCTTGGAGCAGCAGCATGTGGAGATTGCACAGCGGCGATC 607
QY 2841 TGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGACTCCAATGCTATGTGGCTC 2900
Db 608 TGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGACTCCAATGCTATGTGGCTC 667
QY 2901 CTTCCCTTTTGGCCAGTGTATGGAATGTATACATGAGCACTGCCCCCTGAAATTTG 2960
Db 668 CTTCCCTTTTGGCCAGTGTATGGAATGTATACATGAGCACTGCCCCCTGAAATTTG 727
QY 2961 TTCAGGCTACTGTACTCTGTAGTTCATTGCTTGGAGCAACAGGCTGTGCTGCTACTGA 3020
Db 728 TTCAGGCTACTGTACTCTGTAGTTCATTGCTTGGAGCAACAGGCTGTGCTGCTACTGA 787
QY 3021 TCCAGCAATATCTGGCAAGGAAATGCAATAGAGGTTTCTTATAAAGACAGATGAGAT 3080
Db 788 TCCAGCAATATCTGGCAAGGAAATGCAATAGAGGTTTCTTATAAAGACAGATGAGAT 847
```

```
QY 3081 GCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTCTGCTCAATTTCCAGCATGTG 3140
Db 848 GCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTCTGCTCAATTTCCAGCATGTG 907
QY 3141 TCTAGAGGACAGCAGATACAACTGGTCTTTCATTTCACTGTCGAGCTTGCCAAATGCAACGG 3200
Db 908 TCTAGAGGACAGCAGATACAACTGGTCTTTCATTTCACTGTCGAGCTTGCCAAATGCAACGG 967
QY 3201 CCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAGGCAA 3260
Db 968 CCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAGGCAA 1027
QY 3261 GCCTGCGAGACCTGCATATCTGGCTTCTACGCTTACGCTGATCCACCAATGGAGGGAATGTCA 3320
Db 1028 GCCTGCGAGACCTGCATATCTGGCTTCTACGCTGATCCACCAATGGAGGGAATGTCA 1087
QY 3321 GCCATGAAGTGAATGGGACGCGTCTCTGTGCAACACCAACACGCGGCAAGTGTCTTG 3380
Db 1088 GCCATGAAGTGAATGGGACGCGTCTCTGTGCAACACCAACACGCGGCAAGTGTCTTG 1147
QY 3381 CACCACCAAGGGGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCATACCA 3440
Db 1148 CACCACCAAGGGGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCATACCA 1207
QY 3441 AGGAACCCCTCAGAGGACATGTTATTATCTCTTATTGACTATCAGTTCACCTT 3500
Db 1208 AGGAACCCCTCAGAGGACATGTTATTATCTCTTATTGACTATCAGTTCACCTT 1267
QY 3501 TAGTCTATCCAGGAAGATGATCGCTATTACAGAGCTATCAATTTTGTGGTACTCTCTGA 3560
Db 1268 TAGTCTATCCAGGAAGATGATCGCTATTACAGAGCTATCAATTTTGTGGTACTCTCTGA 1327
QY 3561 CGAACAAAACAGGATTTGACATGTTTCATCAATGCTCAAGATTTCAACCTCAACAT 3620
Db 1328 CGAACAAAACAGGATTTGACATGTTTCATCAATGCTCAAGATTTCAACCTCAACAT 1387
QY 3621 CACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTTGTGTTTC 3680
Db 1388 CACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTTGTGTTTC 1447
QY 3681 AAAAAACCAATTAAGAGTACAAAGATGTTTCTCTAATGAGAGTGTGATTTTCGCAA 3740
Db 1448 AAAAAACCAATTAAGAGTACAAAGATGTTTCTCTAATGAGAGTGTGATTTTCGCAA 1507
QY 3741 CCACCCAAATACACTTCTTTGTTTATGTCAGTAAATTTCACTGCGCCATCAAAATCCA 3800
Db 1508 CCACCCAAATACACTTCTTTGTTTATGTCAGTAAATTTCACTGCGCCATCAAAATCCA 1567
QY 3801 GATTGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTCTTCTGTCATTTCTT 3860
Db 1568 GATTGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTCTTCTGTCATTTCTT 1627
QY 3861 CAGTGTGTTCTCTCTTTGCTGCTGGTGGTGTGTTTGGAGATCAAAAGATTTG 3920
Db 1628 CAGTGTGTTCTCTCTTTGCTGCTGGTGGTGTGTTTGGAGATCAAAAGATTTG 1687
QY 3921 TTGGGCTCCAGAGTACAGCAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCC 3980
Db 1688 TTGGGCTCCAGAGTACAGCAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCC 1747
QY 3981 CTTTGCTCTGTAATGTGCTGTTGAAACAGATGAGGAGCCTCTCTGATCTTATTGGGG 4040
Db 1748 CTTTGCTCTGTAATGTGCTGTTGAAACAGATGAGGAGCCTCTCTGATCTTATTGGGG 1807
QY 4041 GAGTATAAGACTGTTCCCAAAACCATTCACCTGGAGCGCTGTTTGGCAACAAAGCCG 4100
Db 1808 GAGTATAAGACTGTTCCCAAAACCATTCACCTGGAGCGCTGTTTGGCAACAAAGCCG 1867
QY 4101 TGTCTCTCTGTGTTGTGAGGCTCCCTCGAGGCTCGGTGGGATCCCTCTCTCTGGGCA 4160
Db 1868 TGTCTCTCTGTGTTGTGAGGCTCCCTCGAGGCTCGGTGGGATCCCTCTCTCTGGGCA 1927
```

```
QY 4161 GTTCAGGTCTTGTGTGGCCAGCGCCTGTGTGGACATTTTCTCAGCAGATGCCGATAGTGA 4220
Db 1928 GTTCAGGTCTTGTGTGGCCAGCGCCTGTGTGGACATTTTCTCAGCAGATGCCGATAGTGA 1987
QY 4221 CAAAGCAGAGTCAGAGCGCTGAGAAACGGAAGCAGCAGCCCTGCACAGCCTGGGAC 4280
Db 1988 CAAAGCAGAGTCAGAGCGCTGAGAAACGGAAGCAGCAGCCCTGCACAGCCTGGGAC 2047
QY 4281 CTGCATCTGA 4290
Db 2048 CTGCATCTGA 2057

RESULT 8
AB011120
LOCUS Homo sapiens mRNA for KIAA0548 protein, partial cds. 5632 bp mRNA linear PRI 10-APR-1998
DEFINITION AB011120
ACCESSION AB011120
VERSION AB011120.1 GI:3043619
KEYWORDS KIAA0548 protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro
JOURNAL DNA Res. 5 (1), 31-39 (1998)
MEDLINE 98290545
PUBMED 9628581
REFERENCE 2 (bases 1 to 5632)
AUTHORS Chara,O., Nagase,T. and Ishikawa,K.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:cdnaif@kazusa.or.jp, Tel:++81-438-52-3913,
Fax:++81-438-52-3914)
FEATURES
Location/Qualifiers
source
1..5632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HH0769"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
gene
1..5632
/gene="KIAA0548"
<1..1361
CDS
/gene="KIAA0548"
/codon_start=3
/product="KIAA0548 protein"
/protein_id="BAA25474.1"
/db_xref="GI:3043620"
/translation="TMSTCPPECSGYCTCSHCLEQPGCGWCTDPSNTGKCKIEGSK
KGPVMPQSAPTGNFYQPQLNSMCLDSRYNWSFIHCPACQNGHSKCTINQSIQCEK
CENLTGKHCETCISGYDPTNGKQPCCKNGHASLQNTNTGKCTTGKVGKDEEC
QICVENRYQGNPLRGTCYITLLIDYQFTSLSQEDDRIYTAINFATPDQNRDLDM
FINAKNPNLNTIWAASFAGTQAGEMPVUVSKTNIKEYKDSFNEKFPDFNHNITP
FVYVSNFTWP:KIQAFSQHSNFDLVQFFTFECSFLSLLVAAVVKIKSCWASR
RREQLRMQOMASRPFASVNVVALETDEEPPDLIGGSIKTVPKFIALEPCFGNKAAL
SVFVRLPRGLGIPPPQSGSLAVASALVDISQWPIVYVKEKSGAVRNKQPPAQPGT
CI"
```

ORIGIN

```
Query Match 31.7%; Score 1361; DB 9; Length 5632;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2930 ATACGATGAGCACCTGCCCCCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCATTGCT 2989
```

1 ATACGATGACACCTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCT 60
2990 TGGAGCAACAGGCTGTGGCTGTGTGTACTGATCCAGCATACTGCGCAAGGGAATGCA 3049
61 TGGAGCAACAGGCTGTGGCTGTGTGTACTGATCCAGCAATACTTGGCAAGGGAATGCA 120
3050 TAGAGGGTTCCTATAAAGGACCACTGCAAGATGCTTCGCAAGCCCTTACAGGAAATTTCT 3109
121 TAGAGGGTTCCTATAAAGGACCACTGCAAGATGCTTCGCAAGCCCTTACAGGAAATTTCT 180
3110 ATCCAGGCCCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGGAGATCAACTGTGCTT 3169
181 ATCCAGGCCCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGGAGATCAACTGTGCTT 240
3170 TCATTCAGTGTCCAGCTTGGCAATGCAACGCGCCACAGTAAATGATCAATCAGAGCATCT 3229
241 TCATTCAGTGTCCAGCTTGGCAATGCAACGCGCCACAGTAAATGATCAATCAGAGCATCT 300
3230 GTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACTGCGATATCTGGCTTCT 3289
301 GTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACTGCGATATCTGGCTTCT 360
3290 ACGTGTATCCCAACCAATGGAGGGAATGTCAGCCATGCAATGGGACGCGCTCTC 3349
361 ACGTGTATCCCAACCAATGGAGGGAATGTCAGCCATGCAATGGGACGCGCTCTC 420
3350 TGTGCAACCAACCAACGCGCAAGTGTCTTCTGACCAACCAAGGGCGTCAAGGGGAGGAGT 3409
421 TGTGCAACCAACCAACGCGCAAGTGTCTTCTGACCAACCAAGGGCGTCAAGGGGAGGAGT 480
3410 GCCAGCTATGTGAGGTGAGAAATGCGATACAGAGAAACCTCTCAGAGGAACATGTTAT 3469
481 GCCAGCTATGTGAGGTGAGAAATGCGATACAGAGAAACCTCTCAGAGGAACATGTTAT 540
3470 ATACTCTTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATT 3529
541 ATACTCTTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATT 600
3530 ACACAGCTATCAATTTTGTGGTACTCTGACGAAACAAACAGGATTTGGACATGTTCA 3589
601 ACACAGCTATCAATTTTGTGGTACTCTGACGAAACAAACAGGATTTGGACATGTTCA 660
3590 TCAATGCTTCAAGAAATTCACCTCAACATCAGTGGGCTGCGATTTCTCAGCTGGAA 3649
661 TCAATGCTTCAAGAAATTCACCTCAACATCAGTGGGCTGCGATTTCTCAGCTGGAA 720
3650 CCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATA 3709
721 CCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATA 780
3710 GTTTCCTAATGAGAGTTGATTTTTCGACACACCCCAATATACATTTCTTTGTTATG 3769
781 GTTTCCTAATGAGAGTTGATTTTTCGACACACCCCAATATACATTTCTTTGTTATG 840
3770 TCAGTAATTTCACTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAATTTTA 3829
841 TCAGTAATTTCACTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAATTTTA 900
3830 TGGACCTGGTACAGTTCTTGTGATCTTTCTTCAGTTGTTCTCTCTTTGCTCTGGTG 3889
901 TGGACCTGGTACAGTTCTTGTGATCTTTCTTCAGTTGTTCTCTCTTTGCTCTGGTG 960
3890 CTGCTGTGTTTGAAGATCAACAAAGTTGTTGGGCTCCAGACGTPAGAGCAACTTC 3949
961 CTGCTGTGTTTGAAGATCAACAAAGTTGTTGGGCTCCAGACGTPAGAGCAACTTC 1020
3950 TTCAGAGATCAACAGATGCCAGCGCTCCCTTTGCTCTGTAAATGTCTGCTTGGAAA 4009
1021 TTCAGAGATCAACAGATGCCAGCGCTCCCTTTGCTCTGTAAATGTCTGCTTGGAAA 1080
4010 CAGATGAGGACCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCCAAAACCCATG 4069

Db 1081 CAGATGAGAGCCCTCTGATCTTATTTGGGGAGATATAAAGACTGTTCCCAAAACCCATTG 1140
Qy 4070 CACTGGAGCCGTTGTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTC 4129
Db 1141 CACTGGAGCCGTTGTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTC 1200
Qy 4130 GAGGCTGGGTGGCATCCCTCTCTGGGCAAGTCAAGTCTGCTGTGCTGAGCCAGCCCTGG 4189
Db 1201 GAGGCTGGGTGGCATCCCTCTCTGGGCAAGTCAAGTCTGCTGTGCTGAGCCAGCCCTGG 1260
Qy 4190 TGGACATTTCTCAGAGATGCCGATAGTGTACAAAGGAGAGTCAAGGAGCCGTGAGAAACC 4249
Db 1261 TGGACATTTCTCAGAGATGCCGATAGTGTGTACAAAGGAGAGTCAAGGAGCCGTGAGAAACC 1320
Qy 4250 GGAAGCAGACGCCCTCTGACAGCCTGGGACCTGCATCTGA 4290
Db 1321 GGAAGCAGACGCCCTCTGACAGCCTGGGACCTGCATCTGA 1361

RESULT 9
AX778545
LOCUS
DEFINITION
Sequence 702 from Patent WO03039443.
ACCESSION
AX778545
VERSION
AX778545.1 GI:32695539
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.
TITLE
Novel genetic markers for leukemias
JOURNAL
Patent: WO 03039443-A 702 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE);
Ludwig-Maximilian-Universitaet Muenchen (DE);
PD Dr. Dr. (DE); Schoch, Claudia (DB); Kern, Wolfgang (DE)
FEATURES
Location/Qualifiers
1..543
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 12.7%; Score 543; DB 6; Length 543;
Best/Local Similarity 100.0%; Pred. No. 2.5e-281; Indels 0; Gaps 0;
Matches 543; Conservative 0; Mismatches 0

Qy 3250 ACCACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCCAATGGA 3309
Db 1 ACCACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCCAATGGA 60
Qy 3310 GGGAAATGTACCCATGCAATGCAATGGGCAACGGCTCTGTGCAACACCAACACGAGG 3369
Db 61 GGGAAATGTACCCATGCAATGCAATGGGCAACGGCTCTGTGCAACACCAACACGAGG 120
Qy 3370 AAGTGTCTTCTGACCAACCAAGGGCTCAAGGGGAGAGTCCAGCTATGTAGGTAGAA 3429
Db 121 AAGTGTCTTCTGACCAACCAAGGGCTCAAGGGGAGAGTCCAGCTATGTAGGTAGAA 180
Qy 3430 AATGATACCAAGGAAACCTCTCAGAGGAACATGTTATATATCTCTTTATTGACTAT 3489
Db 181 AATGATACCAAGGAAACCTCTCAGAGGAACATGTTATATATCTCTTTATTGACTAT 240
Qy 3490 CAGTTTCACTTTAGTCTATCCAGGAGATGATCGCTATTACAGAGTATCAATTTTGTG 3549
Db 241 CAGTTTCACTTTAGTCTATCCAGGAGATGATCGCTATTACAGAGTATCAATTTTGTG 300
Qy 3550 GTACTCTCTGAGCAACAAACAGGAGTTGAGCATGTTTCATCAATGCTCCAGAAATTC 3609
Db 301 GCTACTCTCTGAGCAACAAACAGGAGTTGAGCATGTTTCATCAATGCTCCAGAAATTC 360


```

Db      48926  CCCTGTGTACAGCGGGTGTCTGCAACCTTGGCAGCGGCGGCTGCTGCCCGCGGC 48985
Oy      370    TGGTGGCGGAGCAATGCGACCATGCGGGGGCGGCTTCAG 410
Db      48986  TGGTGGCGGAGCAATGCGACCATGCGGGGGCGGCTTCAG 49026

RESULT 12
AC023420 LOCUS Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT
          SEQUENCE, 15 unordered pieces.
AC023420
VERSION AC023420.4 GI:10280876
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 155869)
          Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
          Bouckhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
          Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
          Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
          Fenesstor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
          Galagan,J., Gardyna,S., Gande,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., Landers,T., Larcocque,K., Lehoczyk,J., Levine,R.,
          Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
          McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
          Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
          Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
          Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
          Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
          Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
          Subramanian,A., Talamas,J., Testave,S., Theodore,J., Tirrell,A.,
          Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
          Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
          Zody,M.
          Direct Submission
          Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          On Sep 23, 2000 this sequence version replaced gi:7767820.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          ----- Genome Center
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WIBR
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence_submissions@genome.wi.mit.edu
          ----- Project Information
          Center project name: L6805
          Center clone name: 737 G 10
          ----- Summary Statistics
          Sequencing vector: M13; M77815; 100% of reads
          Chemistry: Dye-terminator Big Dye; 100% of reads
          Assembly program: Phrap; version 0.960731
          Consensus quality: 145736 bases at least Q40
          Consensus quality: 151534 bases at least Q30
          Consensus quality: 153551 bases at least Q20
          Insert size: 157000; agarose-fp
          Insert size: 154469; sum-cf-ctnigs
          Quality coverage: 4.2 in Q20 bases; agarose-fp
          Quality coverage: 4.3 in Q20 bases; sum-of-ctnigs
          -----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 12284: contig of 12284 bp in length
* 12285 12284: gap of 100 bp
* 12385 12384: contig of 1320 bp in length
* 13705 13804: gap of 100 bp
* 13805 15657: contig of 1853 bp in length
* 15658 15757: gap of 100 bp
* 15758 17755: contig of 1998 bp in length
* 17756 17855: gap of 100 bp
* 17856 19875: contig of 2020 bp in length
* 19876 19875: gap of 100 bp
* 19876 22894: contig of 2919 bp in length
* 22895 22994: gap of 100 bp
* 22995 26992: contig of 3998 bp in length
* 26993 27092: gap of 100 bp
* 27093 31422: contig of 4330 bp in length
* 31423 31522: gap of 100 bp
* 31523 37200: contig of 6338 bp in length
* 37201 38020: gap of 100 bp
* 38021 48644: contig of 10624 bp in length
* 48645 48744: gap of 100 bp
* 48745 69112: contig of 20368 bp in length
* 69113 69212: gap of 100 bp
* 69213 86049: contig of 16337 bp in length
* 86050 86149: gap of 100 bp
* 86150 107477: contig of 21328 bp in length
* 107478 107577: gap of 100 bp
* 107578 131094: contig of 23517 bp in length
* 131095 131194: gap of 100 bp
* 131195 155869: contig of 24675 bp in length.
*
* Location/Qualifiers
*   1..155869
*     /organism="Homo sapiens"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:9606"
*     /chromosome="2"
*     /map="2"
*     /clone="RP11-737G10"
*     /clone_lib="RPC1-11 Human Male BAC"
*   1..12284
*     /note="assembly_fragment"
*     /clone_end:SP6
*     /vector_side:left"
*   12385..13704
*     /note="assembly_fragment"
*   13805..15657
*     /note="assembly_fragment"
*   15758..17755
*     /note="assembly_fragment"
*   17856..19875
*     /note="assembly_fragment"
*   19976..22894
*     /note="assembly_fragment"
*   22995..26992
*     /note="assembly_fragment"
*   27093..31422
*     /note="assembly_fragment"
*   31523..37920
*     /note="assembly_fragment"
*   38021..48644
*     /note="assembly_fragment"
*   48745..69112
*     /note="assembly_fragment"
*   69213..86049
*     /note="assembly_fragment"
*   86150..107477

```


TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 966 21-JAN-2003;
FEATURES Location/Qualifiers
source 1. .246
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.6e-120;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGT 2634
Db 186 CTTGCGAAGATCAATGTGCTCTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTCAGTCTGGATCTGTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTCAGTCTGGATCTGTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 16
AR407079/c
LOCUS AR407079 246 bp DNA linear PAT 19-DEC-2003
DEFINITION Sequence 966 from patent US 6630574.
ACCESSION AR407079
VERSION AR407079.1 GI:40156890
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 246)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mannion, J. and Ran, L.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6630574-A 966 07-OCT-2003;
FEATURES Location/Qualifiers
source 1. .246
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.6e-120;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGT 2634
Db 186 CTTGCGAAGATCAATGTGCTCTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTCAGTCTGGATCTGTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTCAGTCTGGATCTGTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGT 2754

Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 17
AX368256/c
LOCUS AX368256 246 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 966 from Patent WO0204514.
ACCESSION AX368256
VERSION AX368256.1 GI:18856329
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W., Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S., McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0204514-A 966 17-JAN-2002;
FEATURES Location/Qualifiers
source 1. .246
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.6e-120;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGT 2634
Db 186 CTTGCGAAGATCAATGTGCTCTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTCAGTCTGGATCTGTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTCAGTCTGGATCTGTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 18
H009384S15
LOCUS H009384S15 3257 bp DNA linear PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exons 17 and 18.
ACCESSION AF218903
VERSION AF218903.1 GI:8118069
KEYWORDS
SEGMENT 15 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.

Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Horton, L.,
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Klein, J.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE JOURNAL

COMMENT

Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6289165.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3787

Center clone name: 388_K_24

* NOTE: This record contains 125 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 * 857 856: contig of 856 bp in length
 * 857 956: gap of 100 bp
 * 957 1821: contig of 865 bp in length
 * 1822 1821: gap of 100 bp
 * 1822 2776: contig of 855 bp in length
 * 2777 2876: gap of 100 bp
 * 2877 3767: contig of 891 bp in length
 * 3768 3867: gap of 100 bp
 * 3868 4709: contig of 842 bp in length
 * 4710 4809: gap of 100 bp
 * 4810 5655: contig of 846 bp in length
 * 5656 5755: gap of 100 bp
 * 5756 6627: contig of 872 bp in length
 * 6628 6727: gap of 100 bp
 * 6728 7586: contig of 859 bp in length
 * 7587 7686: gap of 100 bp
 * 7687 8241: contig of 855 bp in length
 * 8242 8641: gap of 100 bp
 * 8642 9524: contig of 883 bp in length
 * 9525 9624: gap of 100 bp
 * 9625 10509: contig of 885 bp in length
 * 10510 11474: contig of 865 bp in length
 * 11475 11574: gap of 100 bp
 * 11575 12441: contig of 867 bp in length
 * 12442 12541: gap of 100 bp
 * 12542 13409: contig of 868 bp in length
 * 13410 13509: gap of 100 bp
 * 13510 14439: contig of 930 bp in length
 * 14440 14539: gap of 100 bp
 * 14540 15433: contig of 894 bp in length
 * 15434 15533: gap of 100 bp
 * 15534 16442: contig of 909 bp in length
 * 16443 16542: gap of 100 bp

16543 17421: contig of 879 bp in length
 * 17422 17521: gap of 100 bp
 * 17522 18380: contig of 859 bp in length
 * 18381 18480: gap of 100 bp
 * 18481 19345: contig of 865 bp in length
 * 19346 19445: gap of 100 bp
 * 19446 20327: contig of 882 bp in length
 * 20328 21297: gap of 100 bp
 * 21298 21397: gap of 100 bp
 * 21398 22055: contig of 658 bp in length
 * 22056 22155: gap of 100 bp
 * 22156 23005: contig of 850 bp in length
 * 23006 23105: gap of 100 bp
 * 23106 24016: contig of 911 bp in length
 * 24017 24116: gap of 100 bp
 * 24117 24972: contig of 856 bp in length
 * 24973 25072: gap of 100 bp
 * 25073 25977: contig of 905 bp in length
 * 25978 26077: gap of 100 bp
 * 26078 26952: contig of 875 bp in length
 * 26953 27052: gap of 100 bp
 * 27053 27911: contig of 859 bp in length
 * 27912 28011: gap of 100 bp
 * 28012 28868: contig of 857 bp in length
 * 28869 28968: gap of 100 bp
 * 28969 29624: contig of 656 bp in length
 * 29625 29724: gap of 100 bp
 * 29725 30574: contig of 850 bp in length
 * 30575 30674: gap of 100 bp
 * 30676 31530: contig of 856 bp in length
 * 31531 31630: gap of 100 bp
 * 31631 32491: contig of 861 bp in length
 * 32492 32592: contig of 876 bp in length
 * 32593 33467: contig of 876 bp in length
 * 33468 33567: gap of 100 bp
 * 33568 34337: contig of 870 bp in length
 * 34338 34537: gap of 100 bp
 * 34538 35404: contig of 857 bp in length
 * 35405 35504: gap of 100 bp
 * 35505 36368: contig of 864 bp in length
 * 36369 36468: gap of 100 bp
 * 36469 37353: contig of 885 bp in length
 * 37354 37453: gap of 100 bp
 * 37454 38310: contig of 857 bp in length
 * 38311 38410: gap of 100 bp
 * 38411 39267: contig of 857 bp in length
 * 39268 39367: gap of 100 bp
 * 39368 40217: contig of 850 bp in length
 * 40218 40317: gap of 100 bp
 * 40318 41226: contig of 909 bp in length
 * 41227 41326: gap of 100 bp
 * 41327 42195: contig of 869 bp in length
 * 42196 42295: gap of 100 bp
 * 42296 43180: contig of 885 bp in length
 * 43181 43280: gap of 100 bp
 * 43281 44143: contig of 863 bp in length
 * 44144 44244: gap of 100 bp
 * 44244 45132: contig of 888 bp in length
 * 45132 45231: gap of 100 bp
 * 45232 46149: contig of 918 bp in length
 * 46150 46249: gap of 100 bp
 * 46250 47122: contig of 873 bp in length
 * 47123 47222: gap of 100 bp
 * 47223 48101: contig of 879 bp in length
 * 48102 48201: gap of 100 bp
 * 48202 49041: contig of 840 bp in length
 * 49042 49141: gap of 100 bp
 * 49142 50059: contig of 918 bp in length
 * 50060 50159: gap of 100 bp
 * 50160 51029: contig of 870 bp in length
 * 51030 51129: gap of 100 bp
 * 51130 51995: contig of 866 bp in length

```

* 51996 52095: gap of 100 bp
* 52096 52971: contig of 876 bp in length
* 52972 53071: gap of 100 bp
* 53072 53942: contig of 871 bp in length
* 53943 54042: gap of 100 bp
* 54042 54923: contig of 881 bp in length
* 54923 55023: gap of 100 bp
* 55024 55877: contig of 854 bp in length
* 55878 55977: gap of 100 bp
* 55978 56851: contig of 874 bp in length
* 56852 56951: gap of 100 bp
* 56952 57808: contig of 857 bp in length
* 57809 57909: gap of 100 bp
* 57909 58755: contig of 847 bp in length
* 58756 58855: gap of 100 bp
* 58856 59709: contig of 854 bp in length
* 59710 59809: gap of 100 bp
* 59810 60522: contig of 712 bp in length
* 60522 61501: gap of 100 bp
* 61502 62490: contig of 880 bp in length
* 62491 62590: contig of 889 bp in length
* 62591 63451: contig of 860 bp in length
* 63451 63551: gap of 100 bp
* 63551 64095: contig of 544 bp in length
* 64095 64194: gap of 100 bp
* 64195 65022: contig of 828 bp in length
* 65023 65122: gap of 100 bp
* 65123 65983: contig of 861 bp in length
* 65984 66922: gap of 100 bp
* 66923 67023: contig of 839 bp in length
* 67023 67909: gap of 100 bp
* 67909 68009: contig of 886 bp in length
* 68009 68890: contig of 100 bp
* 68890 68979: gap of 100 bp

Query Match      5.5%; Score 234; DB 2; Length 118540;
Best Local Similarity 100.0%; Pred. No. 2.4e-114;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 CTGAAATGTTTCAGGCTACTGTACTGTAGTCACTGTGAGCAACGAGCTGTGGCT 3010
Db 31011 CTGAAATGTTTCAGGCTACTGTACTGTAGTCACTGTGAGCAACGAGCTGTGGCT 30952

QY 3011 GTGTACTGATCCGACGATATCGCAAGGAAATGATAGAGGGTTCCTATAAGGAC 3070
Db 30951 GGTGTACTGATCCGACGATATCGCAAGGAAATGATAGAGGGTTCCTATAAGGAC 30892

QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGTCAATT 3130
Db 30891 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGTCAATT 30832

QY 3131 CCAGCATGTGTTAGAGGACAGCAGATACAACTGTGTTCTTCACTGTCCAG 3184
Db 30831 CCAGCATGTGTTAGAGGACAGCAGATACAACTGTGTTCTTCACTGTCCAG 30778

RESULT 21
AC055771/c
LOCUS
DEFINITION Homo sapiens chromosome 20 clone RP11-561P16 map 20, WORKING DRAFT
SEQUENCE, 21 unordered pieces.
ACCESSION AC055771
VERSION AC055771.2 GI:7770449
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 167357)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

```

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 167357)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieue,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPeeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,D., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 167357)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieue,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPeeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,D., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 12, 2000 this sequence version replaced gl:7582593.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5007
Center clone name: 561_P_16
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156504 bases at least Q40
Consensus quality: 161679 bases at least Q30
Consensus quality: 163964 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 165357; sum-of-contigs


```
JOURNAL
REFERENCE
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
JOURNAL
Submitted (22-OCT-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
Overlapping Clones: AL132773, AL109804
Cytogenetic Position: 20p
STS Markers: D20S828, D20S813, stSG408, A005005, stS-H22126,
D20S867, D20S752
Additional author information
Haylick, S.J.
Oregon Health Sciences University, Molecular & Medical Genetics
Dept. 3131 SW Sam Jackson Park Road, MC-L103, BSA-4596, Portland OR
97201-3098, USA
phone:1-503-494-6866, fax:1-503-494-6886
e-mail: haylick@ohsu.edu.
FEATURES
source
1. .206647
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p"
/clones="13N6"
ORIGIN
Query Match 5.5%; Score 234; DB 9; Length 206647;
Best Local Similarity 100.0%; Pred. No. 2.3e-114;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2951 CTGAAAATTGTTTCAGGCTACTGTTACCTGTAGTCTGTTGGAGCAACAGGCTGTGGCT 3010
Db 61933 CTGAAAATTGTTTCAGGCTACTGTTACCTGTAGTCTGTTGGAGCAACAGGCTGTGGCT 61992
QY 3011 GGTGTACTGATCCAGCAATCTGCGAAGGAAATGATAGAGGTTCTTATTAAGGAC 3070
Db 61993 GGTGTACTGATCCAGCAATCTGCGAAGGAAATGATAGAGGTTCTTATTAAGGAC 62052
QY 3071 CAGTGAAGATGCTTCCAGAGCCCTACAGAAATTTCTATCCAGCCCTGCTCAATT 3130
Db 62053 CAGTGAAGATGCTTCCAGAGCCCTACAGAAATTTCTATCCAGCCCTGCTCAATT 62112
QY 3131 CCAGCATGTGTCTAGAGGACAGCATACAACTGGTCTTTTCATTCACTGTCCAG 3184
Db 62113 CCAGCATGTGTCTAGAGGACAGCATACAACTGGTCTTTTCATTCACTGTCCAG 62166
RESULT 23
AX326809
LOCUS
DEFINITION
Sequence 5 from Patent WO0178894.
ACCESSION
AX326809
VERSION
AX326809.1 GI:18097525
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Keith, T.
Novel human gene relating to respiratory diseases, obesity, and
inflammatory bowel disease
Patent: WO 0178894-A 5 25-OCT-2001;
Genome Therapeutics Corp. (US)
Location/Qualifiers
1. .207433
/organism="Homo sapiens"
/mol_type="unassigned DNA"
ORIGIN
Query Match 5.2%; Score 223; DB 9; Length 3589;
Best Local Similarity 100.0%; Pred. No. 3.3e-108;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1870 AGCACCATGTATGTTTCGGTGGTTTCAATAGTCTTCCTCCAGGACATCCTGTATTC 1929
Db 1583 AGCACCATGTATGTTTCGGTGGTTTCAATAGTCTTCCTCCAGGACATCCTGTATTC 1642
QY 1930 ACCTCGGAACAGTGTGATCGCATCGAGTACCGCCGTTTGTAGGACGAGCCTGGT 1989
/db_xref="taxon:9606"
ORIGIN
Query Match 5.5%; Score 234; DB 6; Length 207433;
Best Local Similarity 100.0%; Pred. No. 2.3e-114;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2951 CTGAAAATTGTTTCAGGCTACTGTTACCTGTAGTCTGTTGGAGCAACAGGCTGTGGCT 3010
Db 9813 CTGAAAATTGTTTCAGGCTACTGTTACCTGTAGTCTGTTGGAGCAACAGGCTGTGGCT 9872
QY 3011 GGTGTACTGATCCAGCAATCTGCGAAGGAAATGATAGAGGTTCTTATTAAGGAC 3070
Db 9873 GGTGTACTGATCCAGCAATCTGCGAAGGAAATGATAGAGGTTCTTATTAAGGAC 9932
QY 3071 CAGTGAAGATGCTTCCAGAGCCCTACAGAAATTTCTATCCAGCCCTGCTCAATT 3130
Db 9933 CAGTGAAGATGCTTCCAGAGCCCTACAGAAATTTCTATCCAGCCCTGCTCAATT 9992
QY 3131 CCAGCATGTGTCTAGAGGACAGCATACAACTGGTCTTTTCATTCACTGTCCAG 3184
Db 9993 CCAGCATGTGTCTAGAGGACAGCATACAACTGGTCTTTTCATTCACTGTCCAG 10046
RESULT 24
H009384S11
LOCUS
DEFINITION
Homo sapiens attractin precursor (ATRN) gene, exon 12.
ACCESSION
AF218899
VERSION
AF218899.1 GI:8118065
KEYWORDS
11 of 27
SEGMENT
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 3589)
Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
and Duke-Cohan,J.S.
Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
JOURNAL
MEDLINE
20283944
PUBMED
10811918
REFERENCE
2. (bases 1 to 3589)
Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
Schlossman,S.F. and Duke-Cohan,J.S.
Direct Submission
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source
Location/Qualifiers
1. .3589
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
1585..1805
/gene="ATRN"
/number=12
exon
ORIGIN
Query Match 5.2%; Score 223; DB 9; Length 3589;
Best Local Similarity 100.0%; Pred. No. 3.3e-108;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1870 AGCACCATGTATGTTTCGGTGGTTTCAATAGTCTTCCTCCAGGACATCCTGTATTC 1929
Db 1583 AGCACCATGTATGTTTCGGTGGTTTCAATAGTCTTCCTCCAGGACATCCTGTATTC 1642
QY 1930 ACCTCGGAACAGTGTGATCGCATCGAGTACCGCCGTTTGTAGGACGAGCCTGGT 1989
```

Db 1643 ACCTCGACAGTGTGATCGCATCGGAGTGAAGCCGCTTTAGCAGCAGGACCTGGT 1702
 QY 1990 ATTGGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCGTGGCGCTGGCACTGAT 2049
 Db 1703 ATTGGTGTGTGGAACACAGGCTGTCTCAGTGTATCTCGTGGCGCTGGCACTGAT 1762
 QY 2050 GAACAAGAAAGAAAGTAAATCAGATGTTTTTCCAAAGAA 2092
 Db 1763 GAACAAGAAAGAAAGTAAATCAGATGTTTTTCCAAAGAA 1805

RESULT 25
 AC023420/c 155869 bp DNA linear HTG 23-SEP-2000
 LOCUS
 DEFINITION Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT
 SEQUENCE, 15 unordered pieces.

AC023420
 VERSION AC023420.4 GI:10280876
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 2, clone RP11-737G10
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 155869)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beda, F., Boguslavsky, L.,
 Bouckhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
 Choquel, Y., Colangelo, M., Collins, S., Collumore, A., Cooke, P.,
 DeArillano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,
 Feneator, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand, P., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A.,
 Klein, J., Landers, T., Larcocque, K., Lehocsky, J., Levine, R.,
 Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
 McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,
 Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivari, T.M.,
 Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C.,
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A.,
 Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
 Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and
 Zody, M.

Direct Submission
 Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 23, 2000 this sequence version replaced gi:7767820.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L6805
 Center clone name: 737.G.10

Summary Statistics
 Sequencing vector: M13; W7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 145736 bases at least Q40
 Consensus quality: 151534 bases at least Q30
 Consensus quality: 153551 bases at least Q20
 Insert size: 157000; agarose-gel
 Insert size: 154469; sum-of-contigs

Quality coverage: 4.2 in Q20 bases; agarose-gel
 Quality coverage: 4.3 in Q20 bases; sum-of-contigs
 NOTE: This is a 'working draft' sequence. It currently
 consists of 15 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 12284: contig of 12284 bp in length
 12285 12384: gap of 100 bp
 12385 13704: contig of 1320 bp in length
 13705 13804: gap of 100 bp
 13805 15657: contig of 1853 bp in length
 15658 15757: gap of 100 bp
 15758 17755: contig of 1998 bp in length
 17756 17855: gap of 100 bp
 17856 19875: contig of 2020 bp in length
 19876 19975: gap of 100 bp
 19976 22994: contig of 2919 bp in length
 22995 26992: contig of 3998 bp in length
 26993 27092: gap of 100 bp
 27093 31422: contig of 4330 bp in length
 31423 31522: gap of 100 bp
 31523 37920: contig of 6398 bp in length
 37921 38021: gap of 100 bp
 38021 48644: contig of 10624 bp in length
 48645 48744: gap of 100 bp
 48745 69112: contig of 20368 bp in length
 69113 69212: gap of 100 bp
 69213 86049: contig of 16837 bp in length
 86050 86149: gap of 100 bp
 86150 107477: contig of 21328 bp in length
 107478 107577: gap of 100 bp
 107578 131094: contig of 23517 bp in length
 131095 131194: gap of 100 bp
 131195 155869: contig of 24675 bp in length.

Location/Qualifiers
 1..155869
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-737G10"
 /clone_lib="RPC1-11 Human Male BAC"
 1..12284
 /note="assembly_fragment"
 clone_end:sp6
 vector_side:left
 12385..13704
 /note="assembly_fragment"
 13805..15657
 /note="assembly_fragment"
 15758..17755
 /note="assembly_fragment"
 17856..19875
 /note="assembly_fragment"
 19976..22894
 /note="assembly_fragment"
 22995..26992
 /note="assembly_fragment"
 27093..31422
 /note="assembly_fragment"
 31523..37920
 /note="assembly_fragment"
 38021..48644
 /note="assembly_fragment"
 48745..69112
 /note="assembly_fragment"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

```
misc_feature 69213..86049
/note="assembly_fragment"
misc_feature 86150..107477
/note="assembly_fragment"
misc_feature 107578..131094
/note="assembly_fragment"
misc_feature 131195..155869
/note="assembly_fragment"

ORIGIN
Query Match 5.2%; Score 223; DB 2; Length 155869;
Best Local Similarity 100.0%; Pred. No. 2.1e-108; Indels 0; Gaps 0;
Matches 223; Conservative 0; Mismatches 0;

Qy 1870 AGCACCAGTATGTTTCGGTGGTTCAATAGTCTCTCTCAGCGACATCCCTGGTATTC 1929
Db 10396 AGCACCAGTATGTTTCGGTGGTTCAATAGTCTCTCTCAGCGACATCCCTGGTATTC 10337

Qy 1930 ACCTCGGACAGTGTGATCGGCATCGGAGTGAACCGCTGTTTATAGCAGGACCTGGT 1989
Db 10336 ACCTCGGACAGTGTGATCGGCATCGGAGTGAACCGCTGTTTATAGCAGGACCTGGT 10277

Qy 1990 ATTCCGTTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 2049
Db 10276 ATTCCGTTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 10217

Qy 2050 GAACAGAGAAAGTTAAATCAGAATGTTTTCAAAAGAA 2092
Db 10216 GAACAGAGAAAGTTAAATCAGAATGTTTTCAAAAGAA 10174

RESULT 26
H009384514 1728 bp DNA linear PRI 14-DEC-2000
LOCUS H009384514
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 16.
ACCESSION AF218902
VERSION AF218902.1 GI:8118068
KEYWORDS
SEGMENT 14 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1728)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1728)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source Location/Qualifiers
1..1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
926..1169
/gene="ATRN"
/number=16
exon
Query Match 5.2%; Score 221; DB 9; Length 1728;
Best Local Similarity 100.0%; Pred. No. 4.4e-107; Indels 0; Gaps 0;
Matches 221; Conservative 0; Mismatches 0;

Qy 1238 CAACTGGGAATGTGACCAATGATGAGAGTTTTCACATTCATTAATGATGATGGTGT 1297
Db 960 CAACTGGGAATGTGACCAATGATGAGAGTTTTCACATTCATTAATGATGATGGTGT 1019

Qy 1298 TGTGTACCCCTAAGGCAAGGAGCAGTATGTCAGTGGTTGGGCATCTGCACACATTTGA 1357
Db 1020 TGTGTACCCCTAAGGCAAGGAGCAGTATGTCAGTGGTTGGGCATCTGCACACATTTGA 1079

Qy 1358 CACTGAAGAATGCGCGAGTGGTCATGCTGGTCACTCTTTGGTCACTGCCCTCTCTATGAT 1417
Db 1080 CACTGAAGAATGCGCGAGTGGTCATGCTGGTCACTCTTTGGTCACTGCCCTCTCTATGAT 1139
```



```

RESULT 30
H009384S09
LOCUS Homo sapiens attractin precursor (ATRN) gene, exons 9 and 10.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exons 9 and 10.
ACCESSION AF218897
VERSION AF218897.1 GI:8118063
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
1 (bases 1 to 3038)
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
10811918
PUBMED

REFERENCE
2 (bases 1 to 3038)
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
Direct Submission
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

FEATURES
source
1..3038
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/map="20p13"
/number=9
exon
1114..1297
/gene="ATRN"
/number=9
exon
2047..2201
/gene="ATRN"
/number=10
ORIGIN
Query Match 4.3%; Score 186; DB 9; Length 3038;
Best Local Similarity 100.0%; Pred. No. 3.6e-88;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1447 GATAGACACATGGAGTATTATACACACCCAGGGTCCCTTGTGCAAGGGGTACGGC 1506
|||||
Db 1113 GATAGACACATGGAGTATTATACACACCCAGGGTCCCTTGTGCAAGGGGTACGGC 1172
|||||

QY 1507 CATAGCAGTGTTCAGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTACAAGGCT 1566
|||||
Db 1173 CATAGCAGTGTTCAGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTACAAGGCT 1232
|||||

QY 1567 TTCAGTGCATATAGTACCGGCTTGCAGATATCTTACCGATATGATGTGATACCCAG 1626
|||||
Db 1233 TTCAGTGCATATAGTACCGGCTTGCAGATATCTTACCGATATGATGTGATACCCAG 1292
|||||

QY 1627 ATGTGG 1632
|||||
Db 1293 ATGTGG 1298
|||||

RESULT 31
H009384S06
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 6.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 6.
ACCESSION AF218894
VERSION AF218894.1 GI:8118060
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
1 (bases 1 to 1232)
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
10811918
PUBMED

REFERENCE
2 (bases 1 to 1232)
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
Direct Submission
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
20283944
PUBMED

REFERENCE
1 (bases 1 to 1032)
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
Direct Submission
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

FEATURES
Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/map="20p13"
/number=6
exon
529..697
/gene="ATRN"
/number=6
ORIGIN
Query Match 4.0%; Score 173; DB 9; Length 1032;
Best Local Similarity 100.0%; Pred. No. 4.5e-81;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 943 GGTCTGGATGTTTCAGTCTCTACAGCTAACAGTCTTTTGGACTCGAGAGGAATAT 1002
|||||
Db 528 GGTCTGGATGTTTCAGTCTCTACAGCTAACAGTCTTTTGGACTCGAGAGGAATAT 587
|||||

QY 1003 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAACATATATGCG 1062
|||||
Db 588 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAACATATATGCG 647
|||||

QY 1063 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTTCTAGCGTA 1115
|||||
Db 648 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTTCTAGCGTA 700
|||||

RESULT 32
H009384S05
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 5.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 5.
ACCESSION AF218893
VERSION AF218893.1 GI:8118059
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
1 (bases 1 to 1232)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
20283944
PUBMED

REFERENCE
2 (bases 1 to 1232)
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
Direct Submission
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

```


Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 12, 2000 this sequence version replaced gi:7582593.
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: 15007
Center clone name: 561_P_16

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156504 bases at least Q40
Consensus quality: 163679 bases at least Q30
Consensus quality: 163964 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 165357; sum-of-ctg contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-ctg contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1216: contig of 1216 bp in length
* 1217 1316: gap of 100 bp
* 1317 2985: contig of 1669 bp in length
* 2986 3085: gap of 100 bp
* 3086 4628: contig of 1543 bp in length
* 4629 4728: gap of 100 bp
* 4729 7507: contig of 2779 bp in length
* 7508 7607: gap of 100 bp
* 7608 11292: contig of 3685 bp in length
* 11293 11392: gap of 100 bp
* 11393 14213: contig of 2821 bp in length
* 14214 14313: gap of 100 bp
* 14314 17828: contig of 3515 bp in length
* 17829 17928: gap of 100 bp
* 17929 21068: contig of 3140 bp in length
* 21069 21168: gap of 100 bp
* 21169 25035: contig of 3867 bp in length
* 25036 25135: gap of 100 bp
* 25136 28331: contig of 3196 bp in length
* 28332 28431: gap of 100 bp
* 28432 33595: contig of 5164 bp in length
* 33596 33696: gap of 100 bp
* 33696 40346: contig of 6651 bp in length
* 40347 40446: gap of 100 bp
* 40447 43807: contig of 3361 bp in length
* 43808 43907: gap of 100 bp
* 43908 48727: contig of 4820 bp in length
* 48728 48828: gap of 100 bp
* 48828 56183: contig of 7256 bp in length
* 56184 56184: gap of 100 bp
* 56184 67065: contig of 10882 bp in length
* 67066 67165: gap of 100 bp

```

```

* 67166 82881: contig of 15716 bp in length
* 82882 82981: gap of 100 bp
* 82982 101845: contig of 18864 bp in length
* 101846 101945: gap of 100 bp
* 101946 122755: contig of 20810 bp in length
* 122756 122855: gap of 100 bp
* 122856 142457: contig of 19602 bp in length
* 142458 142558: gap of 100 bp
* 142559 167357: contig of 24800 bp in length.

```

FEATURES

source

```

1..167357
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20"
/clone_lib="RP11-561P16"

```

```

/clone="RP11-561P16"

```

```

misc_feature 1..1216
/note="assembly_fragment"
misc_feature 1317..2985
/note="assembly_fragment"
misc_feature 3086..4628
/note="assembly_fragment"
misc_feature 4729..7507
/note="assembly_fragment"
misc_feature 7608..11292
/note="assembly_fragment"
misc_feature 11393..14213
/note="assembly_fragment"
misc_feature 14314..17828
/note="assembly_fragment"
misc_feature 17929..21068
/note="assembly_fragment"
misc_feature 21169..25035
/note="assembly_fragment"
misc_feature 25136..28331
/note="assembly_fragment"
misc_feature 28432..33595
/note="assembly_fragment"
misc_feature 33696..40346
/note="assembly_fragment"
misc_feature 40447..43807
/note="assembly_fragment"
misc_feature 43908..48727
/note="assembly_fragment"
misc_feature 48828..56083
/note="assembly_fragment"
misc_feature 56184..67065
/note="assembly_fragment"
misc_feature 67166..82881
/note="assembly_fragment"
misc_feature 82982..101845
/note="assembly_fragment"
misc_feature 101946..122755
/note="assembly_fragment"
misc_feature 122856..142457
/note="assembly_fragment"
misc_feature 142558..167357
/note="assembly_fragment"

```

ORIGIN

```

Query Match          3.8%   Score 161;   DB 2;   Length 167357;
Best Local Similarity 100.0%;   Pred.No. 7.7e-75;
Matches 161;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 3641 CAGCTGGAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAGGAGT 3700
      |||||
Db 2352 CAGCTGGAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAGGAGT 2411

```

```
QY 3701 ACAAGATAGTTTCTTAATGAGAGTTGATGTTTTCGACACCCCAAAATACACTTTCT 3760
|||
Db 2412 ACAAGATAGTTTCTTAATGAGAGTTGATGTTTTCGACACCCCAAAATACACTTTCT 2471
|||
QY 3761 TTGTTTATGTCAGTAATTCACCTGGCCCAATCAAAATTCAG 3801
|||
Db 2472 TTGTTTATGTCAGTAATTCACCTGGCCCAATCAAAATTCAG 2512
|||

RESULT 35
AC017113 206329 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT
DEFINITION AC017113.4 GI:8570339
ACCESSION HTG; HTGS_PHASE1; HTGS_DRAFT.
VERSION Homo sapiens
KEYWORDS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
SOURCE Waterston,R.H.
ORGANISM 1 (bases 1 to 206329)
REFERENCE 1 (bases 1 to 206329)
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL Waterston,R.H.
REFERENCE 2 (bases 1 to 206329)
AUTHORS The sequence of Homo sapiens clone
TITLE Waterston,R.H.
JOURNAL Direct Submission
AUTHORS Submitted (08-DEC-1999) Genome Sequencing Center, Washington
TITLE University School of Medicine, 444 Forest Park Parkway, St. Louis,
JOURNAL MO 63108, USA
COMMENT On Jun 17, 2000 this sequence version replaced gi:6855220.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0574H07
----- Summary Statistics -----
Sequencing vector: M13; 91%
Sequencing vector: plasmid; 9%
Chemistry: Dye-terminator; 91% of reads
Chemistry: Dye-terminator Big dye; 9% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 184921 bases at least Q40
Consensus quality: 192397 bases at least Q30
Consensus quality: 195862 bases at least Q20
Insert size: 237000; agarose-fp
Insert size: 203329; sum-of-contigs
Quality coverage: 3.84 in Q20 bases; agarose-fp
Quality coverage: 3.66 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1539: contig of 1539 bp in length
* 1540 1639: gap of unknown length
* 1640 3259: contig of 1820 bp in length
* 3260 3359: gap of unknown length
* 3360 5400: contig of 2041 bp in length
* 5401 5500: gap of unknown length
* 5501 7783: contig of 2283 bp in length
* 7784 7883: gap of unknown length
* 7884 11241: contig of 3358 bp in length
* 11242 13441: gap of unknown length
* 13442 13482: contig of 2141 bp in length
*

13483 13582: gap of unknown length
13583 16915: contig of 3333 bp in length
16915 17015: gap of unknown length
17015 21346: contig of 4331 bp in length
21346 21447: gap of unknown length
21447 25103: contig of 3657 bp in length
25103 25204: gap of unknown length
25204 28416: contig of 3212 bp in length
28416 28516: gap of unknown length
28516 32873: contig of 4358 bp in length
32873 36496: contig of 3523 bp in length
36496 36596: gap of unknown length
36596 40649: contig of 4053 bp in length
40649 40749: gap of unknown length
40749 44878: contig of 4129 bp in length
44878 44978: gap of unknown length
44978 50353: contig of 5375 bp in length
50353 55946: contig of 5493 bp in length
55946 56047: gap of unknown length
56047 61421: contig of 5374 bp in length
61421 67819: contig of 6299 bp in length
67819 67919: gap of unknown length
67919 72327: contig of 4408 bp in length
72327 72428: gap of unknown length
72428 77394: gap of unknown length
77394 82626: contig of 5233 bp in length
82626 82727: gap of unknown length
82727 90187: contig of 7460 bp in length
90187 90287: gap of unknown length
90287 96584: contig of 6298 bp in length
96584 106019: contig of 9335 bp in length
106019 113630: contig of 7411 bp in length
113630 113631: gap of unknown length
113631 125151: contig of 11521 bp in length
125151 125252: gap of unknown length
125252 137677: contig of 12426 bp in length
137677 151263: gap of unknown length
151263 151363: contig of 13486 bp in length
151363 165370: contig of 14007 bp in length
165370 165471: gap of unknown length
165471 184705: contig of 19234 bp in length
184705 206329: contig of 21525 bp in length.
184805 206329: contig of 21525 bp in length.

Location/Qualifiers
1..206329
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/clone="RP11-574H7"

FEATURES
source
Query Match 3.8%; Score 161; DB 2; Length 206329;
Best Local Similarity 100.0%; Pred. No. 7.5e-75;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 3641 CAGCTGGAACCCAGGCTGGAGAGAGATGCTGTTTCAAAAACCAACATTAAAGGAGT 3700
Db 173247 CAGCTGGAACCCAGGCTGGAGAGAGATGCTGTTTCAAAAACCAACATTAAAGGAGT 173306
QY 3701 ACAAAGATAGTTTCTTAATGAGAGATGTTTTCGACACCCCAAAATACACTTTCT 3760
Db 173307 ACAAAGATAGTTTCTTAATGAGAGATGTTTTCGACACCCCAAAATACACTTTCT 173366
QY 3761 TTGTTTATGTCAGTAATTCACCTGGCCCAATCAAAATTCAG 3801
|||
```

Db 173367 TTGTTATGTCAGTAATTACCTGGCCATCAAAATTCAG 173407

RESULT 36	AF5311101	LOCUS	AF5311101	4493 bp	mRNA	linear	MAM 19-AUG-2002
DEFINITION	Bos taurus attractin mRNA, complete cds.						
ACCESSION	AF5311101						
VERSION	AF5311101.1		GI:22297307				
KEYWORDS	.						
SOURCE	Bos taurus (cow)						
ORGANISM	Bos taurus						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;						
	Bovidae; Bovinae; Bos.						
REFERENCE	1 (bases 1 to 4493)						

TITLE
JOURNAL
Direct submission
Submitted (18-JUL-2002) Animal sciences, ETH, Tannenstrasse 1.

```

FEATURES
    source
1..4493
    location/Qualifiers
    /organism="Bos taurus"
    /mol_type="mRNA"
    /db_xref="taxon:9913"
    /chromosome="13"
52..4299
    /note="transmembrane form; mahogany protein"
CDS

```

```

/product="attraction"
/protein_id="AAM95445.1"
/db_xref="GI:22297308"
/translations="MVAATARRLRRTVTPAPTRSRNGRCRDWDATRAGRPLGRTGCL
RLPERWRLLLLPLLLLPWAAEAAAASGSSAAEAKECDCRCPVNGGRCNPCT
QCVCPAGWGCQCQCGRGRFLTAPLVAFPGLIVPERDGNITKCTWLIEQGNALIMLR
FQNCPSWDHLYVDGDSIYAPLVAFPGLIVPERDGNITKCTWLIEQGNALIMLR
SDAAAYLNITFNINDEPNCRNGSGCKISNSNTQVCECSNWKGEACDIPHCVN
NCGPFRHGI CNSSDRGVCSQFSEWQPCQSPVPACQFWTRYESYDLKLPRASHKAV
VNGNIMVVGQWFMFNSYMYLVALDLASREWLNRNSVNVYGHSLVALYQDKTY
MYGSKIDTCGNVNELRVLIHNESWLLS PKAKEQYAVUGHSATYLTLANGSVMLV
IFHCPLIYGISVSQBYEDLDKNTVLLSLHTQGNALVQGGYCHSSVYDDRTKALYIHGGYK
IFASANKYRLADDLYVDVDTQWMTLKOSRFPFYLRHTAVTVSGNIMLVFGGNTGNDTSM
SHGAKCSFSDMAYDIDCDRWVSPLKPDHLDVNRFGHSALYNRVTYVFGGNSLLIL
SRLIFPTSEQCAEAGSEAALCAAGCPVRCDMTDSSQCSWELAEQKIHCEKCFSK
RIDPHDRCDLTQCYCTANTNGCWNDCHVPTNHSCTEQIGIFKYDHCPKCDPNMY
YCNKNTSCRSALDQNCQEPNBOICIALPENICGIGHWLVGNSCLKITTARSTYDNA
KLSGRHNHFLASLTQKQVFVFLVQLRIEMQSSQSMKFTLIPVGLRLKNVYWCWE
DMGFFTNHNSLWMPSPSQAQFCSGILSEPTSPRQKAACTINPLNGSVCEPRAHSAKO
CRYPCLATACGCTGGSGECWMCQKVDSDNAYVASFPGQCMWYTMSSCPLEN
CSYCTCSHLEQCGQWCTDTSNTGKGCIEGYSYKGVKPMQSGTNGSYPOPLNLS
MCELDRSYNNWSTIHCPACQCNHGSKCNQISICEKNALTGKHCECTIGSFQYDPIN
GGKQCPQRNCNGHSLQNTWTKCFCTTKGVGDCQCLCEVENRYCGNPLKCTCYVTLL
IDVQFTFESLQEDDRYYTAINVATPDEQNRDLDMFINASQTNLNTIWAASPSAQGTQ
AGEMPVVFNTNIKEYKSFNKEKDFRHNENITFFVTVSNFTPTKIQIAPSOHNF
MDLVAPFSTFCFSGLLIYAAWVKIKQSCWAGRRQEQLRENMQMAARFASVAVV
LETDEBPDLIGGSIKTVKPKIALBPCFNGKAAVLVSFVRLPRGLGIVTTPPGQSGLV
ASALVDISQOMPVYKKGSGAVNRKQOOPGTCI"

```

```
Query Match          3.5%; Score 152; DB 4; Length 4493;
Best Local Similarity 100.0%; Pred. No. 8.7e-70;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

3655	GCTGGAGAGAGATGCCCTGTTGTTTCAAAAACCAATTAAAGGATACAAAGATAGTTTC	3714
QY		
3664	GCTCGAAGAAGAGATGCCCTGTTGTTTCAAAAACCAATTAAAGGATACAAAGATAGTTTC	3723
Db		
3715	TCTAATGAGAAGTTTGATTTTTCGAACCAACCACCAATATCACCTTCTTTGTTTATGTCAGT	3774
QY		
3724	TCTAATGAGAAGTTTGATTTTTCGAACCAACCACCAATATCACCTTCTTTGTTTATGTCAGT	3783
Db		
3775	AATTTCACTGGGCCCATCAAAATTCCAGATTGC	3806
QY		


```

REFERENCE
AUTHORS
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.P. and Duke-Cohan, J.S.
TITLE
JOURNAL
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source
1. .1674
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
810. .947
/gene="ATRN"
/number=19
exon
ORIGIN
Query Match 3.3%; Score 141; DB 9; Length 1674;
Best Local Similarity 100.0%; Pred. No. 8.8e-64;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3182 CAGTTCCTCCATGCAAGCGGCACAGTAATGCATCAATCAGACCATCTGTGAGAGTGTG 3241
Db 807 CAGTTCCTCCATGCAAGCGGCACAGTAATGCATCAATCAGACCATCTGTGAGAGTGTG 866
Qy 3242 AGAACCTTGACACAGGCAAGCAGCTCGGAGACCTGCATATCTGGTTCTACGGTGATCCCA 3301
Db 867 AGAACCTTGACACAGGCAAGCAGCTCGGAGACCTGCATATCTGGTTCTACGGTGATCCCA 926
Qy 3302 CCAATGGAGGAAATGTCAGC 3322
Db 927 CCAATGGAGGAAATGTCAGC 947
RESULT 39
AC013324
LOCUS AC013324 118540 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-388K24 map 2, LOW-PASS
SEQUENCE SAMPLING.
AC013324
VERSION AC013324.3 GI:9123914
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Biren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 2, clone RP11-388K24
Unpublished
2 (bases 1 to 118540)
Biren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouckgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Horton, L.,
Jalagani, J., Kardana, S., Grant, G., Hegos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6289165.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3787
Center clone name: 388_K_24
-----

```

```

* NOTE: This record contains 125 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

1
856: contig of 856 bp in length
857
956: gap of 100 bp
957
1821: contig of 863 bp in length
1822
1821: gap of 100 bp
1922
2776: contig of 855 bp in length
2777
2876: gap of 100 bp
2877
3767: contig of 891 bp in length
3768
3867: gap of 100 bp
3868
4709: contig of 842 bp in length
4710
4809: gap of 100 bp
4810
5855: contig of 846 bp in length
5856
5755: gap of 100 bp
5857
6627: contig of 872 bp in length
6628
6727: gap of 100 bp
6728
7586: contig of 859 bp in length
7587
7686: gap of 100 bp
7688
8841: contig of 855 bp in length
8842
8641: gap of 100 bp
8843
9524: contig of 883 bp in length
9525
9624: gap of 100 bp
9625
10510
10609: contig of 885 bp in length
10511
11474: contig of 865 bp in length
11475
11574: gap of 100 bp
12441: contig of 867 bp in length
12442
12541: gap of 100 bp
12542
13409: contig of 868 bp in length
13410
13509: gap of 100 bp
13510
14439: contig of 930 bp in length
14440
14539: gap of 100 bp
14540
15433: contig of 894 bp in length
15434
15533: gap of 100 bp
15534
16442: contig of 909 bp in length
16443
16542: gap of 100 bp
16543
17421: contig of 879 bp in length
17422
17521: gap of 100 bp
17522
18380: contig of 859 bp in length
18381
18480: gap of 100 bp
18481
19345: contig of 865 bp in length
19346
19445: gap of 100 bp
19446
20327: contig of 882 bp in length
20328
20427: gap of 100 bp
20428
21297: contig of 870 bp in length
21298
21397: gap of 100 bp
22055: contig of 658 bp in length
22056
22155: gap of 100 bp
22156
23005: contig of 850 bp in length
23006
23105: gap of 100 bp
23106
24016: contig of 911 bp in length
24017
24116: gap of 100 bp
24117
24972: contig of 856 bp in length
24973
25072: gap of 100 bp

```

```

* 25073 25977: contig of 905 bp in length
* 25978 26077: gap of 100 bp
* 26078 26952: contig of 875 bp in length
* 26953 27052: gap of 100 bp
* 27053 27911: contig of 859 bp in length
* 27912 28011: gap of 100 bp
* 28012 28868: contig of 857 bp in length
* 28869 28968: gap of 100 bp
* 28969 29224: contig of 656 bp in length
* 29225 29724: gap of 100 bp
* 29725 30574: contig of 850 bp in length
* 30575 30674: gap of 100 bp
* 30675 31530: contig of 856 bp in length
* 31531 31630: gap of 100 bp
* 31631 32491: contig of 861 bp in length
* 32492 32591: gap of 100 bp
* 32592 33467: contig of 876 bp in length
* 33468 33567: gap of 100 bp
* 33568 34437: contig of 870 bp in length
* 34438 34537: gap of 100 bp
* 34538 35404: contig of 867 bp in length
* 35405 35504: gap of 100 bp
* 35505 36368: contig of 864 bp in length
* 36369 36468: gap of 100 bp
* 36469 37353: contig of 885 bp in length
* 37354 37453: gap of 100 bp
* 37454 38310: contig of 857 bp in length
* 38311 38410: gap of 100 bp
* 38411 39267: contig of 857 bp in length
* 39268 39367: gap of 100 bp
* 39368 40217: contig of 850 bp in length
* 40218 40317: gap of 100 bp
* 40318 41226: contig of 903 bp in length
* 41227 41326: gap of 100 bp
* 41327 42195: contig of 869 bp in length
* 42196 42295: gap of 100 bp
* 42296 43180: contig of 885 bp in length
* 43181 43280: gap of 100 bp
* 43281 44143: contig of 863 bp in length
* 44144 44243: gap of 100 bp
* 44244 45131: contig of 888 bp in length
* 45132 45231: gap of 100 bp
* 45232 46149: contig of 918 bp in length
* 46150 46249: gap of 100 bp
* 46250 47122: contig of 873 bp in length
* 47123 47222: gap of 100 bp
* 47223 48101: contig of 879 bp in length
* 48102 48201: gap of 100 bp
* 48202 49041: contig of 840 bp in length
* 49042 49141: gap of 100 bp
* 49142 50059: contig of 918 bp in length
* 50060 50159: gap of 100 bp
* 50160 51029: contig of 870 bp in length
* 51030 51129: gap of 100 bp
* 51130 51995: contig of 866 bp in length
* 51996 52095: gap of 100 bp
* 52096 52971: contig of 876 bp in length
* 52972 53071: gap of 100 bp
* 53072 53942: contig of 871 bp in length
* 53943 54042: gap of 100 bp
* 54043 54923: contig of 881 bp in length
* 54924 55023: gap of 100 bp
* 55024 55877: contig of 854 bp in length
* 55878 55977: gap of 100 bp
* 55978 56851: contig of 874 bp in length
* 56852 56951: gap of 100 bp
* 56952 57808: contig of 857 bp in length
* 57809 57908: gap of 100 bp
* 57909 58755: contig of 847 bp in length
* 58756 58855: gap of 100 bp
* 58856 59709: contig of 854 bp in length
* 59710 59809: gap of 100 bp
* 59810 60521: contig of 712 bp in length

```

```

* 60522 60421: gap of 100 bp
* 60422 61501: contig of 880 bp in length
* 61502 61602: gap of 100 bp
* 61603 62490: contig of 889 bp in length
* 62491 62590: gap of 100 bp
* 62591 63450: contig of 860 bp in length
* 63451 63550: gap of 100 bp
* 63551 64094: contig of 544 bp in length
* 64095 64194: gap of 100 bp
* 64195 65022: contig of 828 bp in length
* 65023 65122: gap of 100 bp
* 65123 65983: contig of 861 bp in length
* 65984 66083: gap of 100 bp
* 66084 66922: contig of 839 bp in length
* 66923 67022: gap of 100 bp
* 67023 67908: contig of 885 bp in length
* 67909 68008: gap of 100 bp
* 68009 68879: contig of 871 bp in length
* 68880 68979: gap of 100 bp

Query Match 3.2%; Score 136; DB 2; Length 118540;
Best Local Similarity 99.5%; Pred. No. 2.7e-61;
Matches 186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2764 GCAACACAGTCTAAGCAGTCCGCGACACCATGTGCTTGAGACAGCATGTGGAGAT 2823
Db 78072 GCAACACAGTCTAAGCAGTCCGCGACACCATGTGCTTGAGACAGCATGTGGAGAT 78131
QY 2824 TGCACACAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAACAGTGTGTGACTCC 2883
Db 78132 TGCACACAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAACAGTGTGTGACTCC 78191
QY 2884 AATGCTATGTGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTATACATGAGCACC 2943
Db 78192 AATGCTATGTGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTATACATGAGCACC 78251
QY 2944 TGCCCCC 2950
Db 78252 TGCCCCC 78258

RESULT 40
H009384S04
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 4.
DEFINITION AF218892
ACCESSION AF218892.1 GI:8118058
VERSION AF218892.1
KEYWORDS 4 of 27
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1152)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.P., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1152)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.P., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submision
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source
Location/Qualifiers
1..1152
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```



```

repeat_region 11683..11848
/Note="L1 repeat: matches 3736..3898 of consensus"
repeat_region 11849..12157
/Note="AluSg repeat: matches 1..309 of consensus"
repeat_region 12158..12182
/Note="L1 repeat: matches 3712..3736 of consensus"
repeat_region 12200..12396
/Note="L1 repeat: matches 1702..1912 of consensus"
repeat_region 12417..12586
/Note="L1 repeat: matches 7758..7913 of consensus"
repeat_region 12634..12914
/Note="AluX repeat: matches 6..312 of consensus"
repeat_region 12958..13404
/Note="L1 repeat: matches 1267..1711 of consensus"
repeat_region 13486..13733
/Note="AluSg1 repeat: matches 36..285 of consensus"
repeat_region 14313..14415
/Note="MER86 repeat: matches 89..196 of consensus"
repeat_region 14493..15186
/Note="L1 repeat: matches 4409..5090 of consensus"
repeat_region 15187..15487
/Note="AluX repeat: matches 12..312 of consensus"
repeat_region 15488..15539
/Note="L1 repeat: matches 4358..4409 of consensus"
misc_feature 15537..15942
/Note="match: STS: Em:R15876"
misc_feature 15537..15748
/Note="match: STS: Em:G15134"
gene complement(15767..17243)
/Note="dJ581P3.1"
CDS
complement(15767..17243)
/Note="dJ581P3.1"
60kD (SAP61) pseudogene
match: CDNA: Em:X81789 Em:U08815
match: EST: Em:A112866 Em:T56818 Em:AW470436 Em:AA300516
Em:AA703023 Em:AA760210 Em:AA054241 Em:AA443032 Em:R17811
Em:AW493824 Em:H67476 Em:A1020542 Em:AA343902 Em:H89818
Em:AA308239 Em:AA311786 Em:A1541046 Em:AA239392
Em:AA694306 Em:AA086662 Em:R15877 Em:A1226265 Em:P22705
Em:A1155866 Em:C06258 Em:AA086076 Em:AA081903 Em:A1019200
Em:AV138764 Em:AA06960 Em:A1020796 Em:R00790 Em:AA308500
Em:AA003841 Em:AV290458 Em:AV166585 Em:AA074650
Em:AA135375 Em:AA355911 Em:AA730707 Em:AA068453
Em:AA306883 Em:AL035731 Em:AA992803 Em:AL035732
Em:AA854771 Em:AA320502 Em:AA599259 Em:A1374964
Em:A1338280 Em:AA662262 Em:AA642724 Em:W83741 Em:AA662366
Em:AA463475 Em:W35845 Em:A1269312 Em:AL118208 Em:AA986711
Em:A1374716 Em:A1908435 Em:C82711 Em:AL039268 Em:AW246145
match: proteins: Tr:Q12874 Wp:FLH5 Tr:O59706 Tr:Q22469
Tr:O46106 Sw:P19736"
/pseudo
/codon_start=1
/evidence=not_experimental
16008..16100
/Note="match: GSS: Em:AQ279830"
misc_feature 16104..16193
/Note="match: GSS: Em:AQ074332"
misc_feature complement(16881..16954)
/Note="match: GSS: Em:AQ975997"
/pseudo
16980..17146
/Note="match: GSS: Em:AQ373489"
misc_feature 17150..17265
/Note="match: GSS: Em:AQ373489"
repeat_region 17389..17749
/Note="L2 repeat: matches 1140..1536 of consensus"
repeat_region 17859..18030
/Note="L2 repeat: matches 1791..1990 of consensus"
repeat_region 18405..18442
/Note="L1 repeat: matches 19 copies 2 mer gt 97% conserved"
misc_feature complement(19050..19224)
/Note="match: GSS: Em:AQ472547"
repeat_region 19051..19092
/Note="21 copies 2 mer ag 90% conserved"
misc_feature complement(19093..19250)
/Note="match: GSS: Em:AQ271994"
repeat_region 19142..19175
Query Match 3.1%; Score 131; DB 9; Length 73666;
Best Local Similarity 100.0%; Pred. No. 1.5e-58;
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 607 AGTGGCTCATTTGTTCTGAGAGATGCAATGAGATGTCCCTGAGGTGTGTTCCACA 666
DB 71998 AGTGGCTCATTTGTTCTGAGAGATGCAATGAGATGTCCCTGAGGTGTGTTCCACA 72057
QY 667 TAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTATTAATTTGACTGGATTAAAT 726
DB 72058 TAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTATTAATTTGACTGGATTAAAT 72117
QY 727 ATTACTTACAG 737
DB 72118 ATTACTTACAG 72128
RESULT 42
H009384S27 5790 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 30 and complete
DEFINITION cds, alternatively spliced.
ACCESSION AF218915.1 GI:8118081
VERSION AF218915.1
KEYWORDS 27 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 5790)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schloeman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 5790)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schloeman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
Location/Qualifiers
1..5790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="20p13"
/order(AF218899.1:745..1576,AF218890.1:1..3711,
AF218891.1:1..3058,AF218892.1:1..1152,AF218893.1:1..1232,
AF218894.1:1..1032,AF218895.1:1..1575,AF218896.1:1..2018,
AF218897.1:1..3038,AF218898.1:1..1552,AF218899.1:1..3589,
AF218900.1:1..1267,AF218901.1..3112,AF218902.1:1..1728,
AF218903.1:1..3257,AF218904.1:1..1674,AF218905.1:1..1675,
AF218906.2:1..966,AF218907.1:1..1781,AF218908.1:1..1704,
AF218909.1:1..1976,AF218910.1:1..1364,AF218911.1:1..2214,
AF218912.1:1..1556,AF218913.1:1..1592,AF218914.1:1..1394,
1..5060)
/gene="ATRN"
join(AF218899.1:745..1233,AF218890.1:2877..2960,
AF218891.1:810..923,AF218892.1:818..946,

```

```
AF218893.1:621..826,AF218894.1:529..697,
AF218895.1:11343..1433,AF218896.1:1926..1169,
AF218897.1:11114..1297,AF218898.1:2047..2201,
AF218899.1:1659..743,AF218899.1:1585..1805,
AF218900.1:1728..849,AF218901.1:440..578,
AF218901.1:2163..2353,AF218902.1:533..752,
AF218903.1:11441..1626,AF218903.1:2189..2422,
AF218904.1:1810..947,AF218905.1:833..926,
AF218906.2:49..101,AF218907.1:802..896,
AF218908.1:873..951,AF218909.1:392..549,
AF218911.1:1348..1409,AF218912.1:748..826,
AF218913.1:1747..854,AF218914.1:735..849,660..5060)
/gene="ATRN"
/product="membrane attractin precursor"
join(AF218899.1:824..1233,AF218899.1:2877..2960,
AF218899.1:1810..923,AF218899.1:1818..946,
AF218899.1:1621..826,AF218899.1:529..697,
AF218899.1:11343..1433,AF218899.1:1926..1169,
AF218899.1:11114..1297,AF218899.1:2047..2201,
AF218899.1:1659..743,AF218899.1:1585..1805,
AF218900.1:1728..849,AF218901.1:440..578,
AF218901.1:2163..2353,AF218902.1:533..752,
AF218903.1:11441..1626,AF218903.1:2189..2422,
AF218904.1:1810..947,AF218905.1:833..926,
AF218906.2:49..101,AF218907.1:802..896,
AF218908.1:873..951,AF218909.1:392..549,
AF218911.1:1348..1409,AF218912.1:748..826,
AF218913.1:1747..854,AF218914.1:735..849,660..784)
/gene="ATRN"
/note="Method: conceptual translation with partial peptide
sequencing"
/codon_start=1
/product="membrane attractin precursor"
/protein_id="AA72881.1"
/db_xref="GI:8118082"
/translation="MVAAATAEARLRRTAATAAAGRGSGPHWDVTRAGRGPLG
AGLEPLRLSPLEPLRLGLLLPPLLLPCEAAASAAAASVGSAAAEKED
RPGVNGCPNGPQQCPAGVGEQCHGCRFRFGSSGFVTDGFGNYKTKCTM
LIEQPNRINELRNFATECSDEHLYVDGDSIYAPLVAFLGIYPERDGNETVE
VWATGVALHFFSDAAYLTGFNITYFDMCPNCSGRGECISNDVCECGEN
WKGACDIPHTDNCGRPHGICNSDVRCSCFSDWQPGQCSVPANQFVREY
SNLKLPRASHKAVNGIMVGVGMFNHSDYNNVLYDLASREWLPLNRSVNVVR
YHSLALYKOKIYMGKIDSTGVNINELRVFHLHNSWVLLTPKAKEAVVGHSAH
ITVLKNGRVMLVIFGHCPLYGISNVEYDLQNTWLSILHQALVQGGYSHSYVD
HRTALYVHGKYAFSANKYRLADLYRDYDTQMOTILKDSFRFLYHTAVISGTM
LVFGNTDNTSMGSHAKCFSSDFMAYDIACDMSVLPRLDHDVNRFGHSAVLNS
TMVYFGFNSLLIDILVFTSEOCDAHRSAACIAGPGRVCVNTGSSOCI SWALAT
DEQEKLSCEPSEKXETLDHBCDQHTCYCTANTNDCHNCHCVPRNCSCEGLS
IFRIENCPKNMPTIKNKTSCRSCLADQNCQEPNRCQECALPENICGIGWLVGNS
CLKITAKENYDNLKIFRNHALLSLTTQKRFVFLKOLRINQSSQSKLITPLW
VGLRKINYSWCDMSDFPNTSLQWPESEPDAGFCGLSEPTSLKKAATCINFLN
GSVCERPAHSAKQCTPCLARTACGCTSGSSECMMCSNNKQCVDSNAYVAFPPFQ
CMWYWTSTCPENCSCVCTCHLEQPGCWCTDPSTGTGKICIGSYKGVKMPQ
APTGNFVOPPLNSSLMLDSRYNNSFIHCPACQNGHSCINQICEKCNLTGKH
CETCISFGDPTNGGQKCKNGHASLQNTWTGKCFCTTKGVKGDCECOLCEVRY
QGNELRGCTYLLIDVQFTPSLCKDDRYITAINFVATPEDEDRDLDMFNASKNPN
LNIITWASFSAGTQAGBEMFVYKTNKEYDSFNEKFDNRHNPITFFVYSNFTW
PIKQAFASHSNFMIDVQFTPTFFSCFLSLVAAYVWKIKOSCWASRRERQLREM
QOMASRPFASVDEVPDDEPDLIGSITKVPKPIALEPFCGNKAVLSVFLVLRPG
LGGTPPQOSGLAVASLVDISQMPVYKESGAVNRKQPPAQPGTCI"
join(AF218899.1:1073..1233,AF218899.1:2877..2960,
AF218899.1:1810..923,AF218899.1:1818..946,
AF218899.1:1621..826,AF218899.1:529..697,
AF218899.1:11343..1433,AF218899.1:1926..1169,
AF218899.1:11114..1297,AF218899.1:2047..2201,
AF218899.1:1659..743,AF218899.1:1585..1805,
AF218900.1:1728..849,AF218901.1:440..578,
AF218901.1:2163..2353,AF218902.1:533..752,
AF218903.1:11441..1626,AF218903.1:2189..2422,
AF218904.1:1810..947,AF218905.1:833..926,
AF218906.2:49..101,AF218907.1:802..896,
AF218908.1:873..951,AF218909.1:392..549,
AF218911.1:1348..1409,AF218912.1:748..826,
AF218913.1:1747..854,AF218914.1:735..849,660..784)
```

mat_peptide

```
AF218913.1:1747..854,AF218914.1:735..849,660..781)
/gene="ATRN"
/product="membrane attractin"
660..5060
/gene="ATRN"
/number=30
polyA_signal
5038..5043
/gene="ATRN"
ORIGIN
Query Match 3.0%; Score 127; DB 9; Length 5790;
Best Local Similarity 100.0%; Pred. No. 2.9e-56;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4164 AGCTCTTGCTGGCCAGCGCCCTGTGTGACATTTCTAGCAGATGCCATAGTGTACAA 4223
Db 658 AGCTCTTGCTGGCCAGCGCCCTGTGTGACATTTCTAGCAGATGCCATAGTGTACAA 717
QY 4224 GGAGAAAGTCAGGAGCGCTGAGAAACCGGAAGCAGAGAGCCCTTGCACAGCTGGGACCTG 4283
Db 718 GGAGAAAGTCAGGAGCGCTGAGAAACCGGAAGCAGAGAGCCCTTGCACAGCTGGGACCTG 777
QY 4284 CATCTCA 4290
Db 778 CATCTCA 784
RESULT 43
AC017113/c
LOCUS Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT
DEFINITION AC017113
ACCESSION AC017113.4 GI:8570339
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206329)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 206329)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jun 17, 2000 this sequence version replaced gi:6855220.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0574H07
----- Summary Statistics -----
Sequencing vector: M13; 91%
Sequencing vector: plasmid; 9%
Chemistry: Dye-terminator Big Dye; 9% of reads
Chemistry: Dye-terminator Big Dye; 9% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 184921 bases at least Q40
Consensus quality: 192397 bases at least Q30
Consensus quality: 195862 bases at least Q20
Insert size: 237000; agarose-fp
Insert size: 203329; sum-of-ctnigs
Quality coverage: 3.84 in Q20 bases; agarose-fp
Quality coverage: 3.66 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
consists of 31 contigs. The true order of the pieces
```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1539: contig of 1539 bp in length
* 1540 1639: gap of unknown length
* 3259: contig of 1620 bp in length
* 3359: gap of unknown length
* 3360 5400: contig of 2041 bp in length
* 5401 5500: gap of unknown length
* 7783: contig of 2283 bp in length
* 7784 7883: gap of unknown length
* 7884 11241: contig of 3358 bp in length
* 11242 11341: gap of unknown length
* 11342 13482: contig of 2141 bp in length
* 13483 13582: gap of unknown length
* 13583 16915: contig of 3333 bp in length
* 16916 17015: gap of unknown length
* 21346: contig of 4331 bp in length
* 21347 25103: gap of unknown length
* 25104 25203: gap of unknown length
* 25204 28415: contig of 3212 bp in length
* 28416 32873: contig of 4358 bp in length
* 32874 32973: gap of unknown length
* 32974 36496: contig of 3523 bp in length
* 36497 36596: gap of unknown length
* 36597 40649: contig of 4053 bp in length
* 40650 40749: gap of unknown length
* 40750 44878: contig of 4129 bp in length
* 44879 44978: gap of unknown length
* 44979 50353: contig of 5375 bp in length
* 50354 50453: gap of unknown length
* 50454 55946: contig of 5493 bp in length
* 55947 56046: gap of unknown length
* 56047 61420: contig of 5374 bp in length
* 61421 61520: gap of unknown length
* 61521 67819: contig of 6299 bp in length
* 67820 67919: gap of unknown length
* 67920 72327: contig of 4408 bp in length
* 72328 72427: gap of unknown length
* 72428 77293: contig of 4866 bp in length
* 77294 77393: gap of unknown length
* 77394 82626: contig of 5233 bp in length
* 82627 82726: gap of unknown length
* 82727 90186: contig of 7460 bp in length
* 90187 90286: gap of unknown length
* 90287 96584: contig of 6298 bp in length
* 96585 96684: gap of unknown length
* 96685 106019: contig of 9335 bp in length
* 106020 106119: gap of unknown length
* 106120 113330: contig of 7411 bp in length
* 113331 113630: gap of unknown length
* 113631 125151: contig of 11521 bp in length
* 125152 125251: gap of unknown length
* 125252 137677: contig of 12426 bp in length
* 137678 137777: gap of unknown length
* 137778 151263: contig of 13486 bp in length
* 151264 151363: gap of unknown length
* 151364 165370: contig of 14007 bp in length
* 165371 165470: gap of unknown length
* 165471 184704: contig of 19234 bp in length
* 184705 184804: gap of unknown length
* 206329: contig of 21525 bp in length.

FEATURES
source

Location/Qualifiers
1..206329
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"

ORIGIN
/clone="RP11-574H7"
Query Match 3.0%; Score 127; DB 2; Length 206329;
Best Local Similarity 100.0%; Pred. No. 1.9e-56; Indels 0; Gaps 0;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4164 AGGTCTTGTGTGGCAGCGCCCTGTGACATTTCTCAGCAGATGCCGATAGTACAA 4223
Db 159328 AGGTCTTGTGTGGCAGCGCCCTGTGACATTTCTCAGCAGATGCCGATAGTACAA 159269
QY 4224 GGAGAGTCAGGAGCGGTGAGAAACCGGAAGCAGCAGCCCTGCACAGCTGGACCTG 4283
Db 159268 GGAGAGTCAGGAGCGGTGAGAAACCGGAAGCAGCAGCCCTGCACAGCTGGACCTG 159209
QY 4284 CATCTGA 4290
Db 159208 CATCTGA 159202
RESULT 44
H009384S26 1394 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 29.
DEFINITION AF218914
ACCESSION AF218914
VERSION AF218914.1 GI:8118080
KEYWORDS 26 of 27
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
and Duke-Cohan,J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1394)
AUTHORS Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
Schlossman,S.F. and Duke-Cohan,J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
Location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
735..849
/gene="ATRN"
/number=29
exon
ORIGIN
Query Match 2.8%; Score 119; DB 9; Length 1394;
Best Local Similarity 100.0%; Pred. No. 7.3e-52; Indels 0; Gaps 0;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4049 AGACTGTTCCCAAAACCCATTGACCTGGAGCCCTGTTTGGCAACAAACCGCTGCTCT 4108
Db 733 AGACTGTTCCCAAAACCCATTGACCTGGAGCCCTGTTTGGCAACAAACCGCTGCTCT 792
QY 4109 CTGTGTTTGTGAGGCTCCCTCCGAGCCCTGGGTCATCCCTCTCGGAGTCAGGT 4167
Db 793 CTGTGTTTGTGAGGCTCCCTCCGAGCCCTGGGTCATCCCTCTCGGAGTCAGGT 851
RESULT 45


```
H009384S03
LOCUS      H009384S03      3058 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 3.
ACCESSION AF218891
VERSION    AF218891.1 GI:8118057
KEYWORDS   3 of 27
SEGMENT    Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 3058)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE      Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE    20283944
PUBMED     10811918
REFERENCE  2 (bases 1 to 3058)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE      Direct Submission
JOURNAL    Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES   Location/Qualifiers
source     1..3058
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20p13"
            810..923
            /gene="ATRN"
            /number=3
exon       495 GCCTAATAGATATAGAGACTCGTTTCATCATTCATTGCTACAGAGTGTAGTGGACCA 554
            |||
            810 GCCTAATAGATATAGAGACTCGTTTCATCATTCATTGCTACAGAGTGTAGTGGACCA 869
            |||
            555 TTATATGTTTATGATGGGACTCAATTATGACCGCTAGTCTGCATTAG 608
            |||
            870 TTATATGTTTATGATGGGACTCAATTATGACCGCTAGTCTGCATTAG 923
            |||
RESULT 46
AF273610
LOCUS      AF273610      237 bp      mRNA      linear      MAM 09-JUL-2000
DEFINITION Sus scrofa mahogany mRNA, partial cds.
ACCESSION AF273610
VERSION    AF273610.1 GI:8980838
KEYWORDS   2 (bases 1 to 237)
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
REFERENCE  1 (bases 1 to 237)
AUTHORS   Matteri, R.L. and Dyer, C.J.
TITLE      Sus scrofa mahogany mRNA
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 237)
AUTHORS   Matteri, R.L. and Dyer, C.J.
TITLE      Direct Submission
JOURNAL    Submitted (01-JUN-2000) Animal Physiology Research Unit, USDA -
Agricultural Research Service, University of Missouri, 920 East
Campus Dr., Room S-143 ASRC, Columbia, MO 65211, USA
FEATURES   Location/Qualifiers
source     1..1592
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20p13"
            747..854
            /gene="ATRN"
            /number=28
exon       2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 2304
            |||
            125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 184
            |||
            2305 CAGAACTGCAGTGGGAGCCCGGATCAGGATCAGGATGCGCTGCCGAAA 2357
            |||
            185 CAGAACTGCAGTGGGAGCCCGGATCAGGATGCGCTGCCGAAA 237
            |||
Query Match 2.6%; Score 113; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.6e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 2304
Db 125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 184
QY 2305 CAGAACTGCAGTGGGAGCCCGGATCAGGATCAGGATGCGCTGCCGAAA 2357
Db 185 CAGAACTGCAGTGGGAGCCCGGATCAGGATGCGCTGCCGAAA 237
RESULT 47
H009384S25
LOCUS      H009384S25      1592 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 28.
ACCESSION AF218913
VERSION    AF218913.1 GI:8118079
KEYWORDS   25 of 27
SEGMENT    Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1592)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE      Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE    20283944
PUBMED     10811918
REFERENCE  2 (bases 1 to 1592)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE      Direct Submission
JOURNAL    Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES   Location/Qualifiers
source     1..1592
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20p13"
            747..854
            /gene="ATRN"
            /number=28
exon       2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 2304
            |||
            125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 184
            |||
            2305 CAGAACTGCAGTGGGAGCCCGGATCAGGATCAGGATGCGCTGCCGAAA 2357
            |||
            185 CAGAACTGCAGTGGGAGCCCGGATCAGGATGCGCTGCCGAAA 237
            |||
Query Match 2.6%; Score 110; DB 9; Length 1592;
Best Local Similarity 100.0%; Pred. No. 5.4e-47;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
H009384S03
LOCUS      H009384S03      3058 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 3.
ACCESSION AF218891
VERSION    AF218891.1 GI:8118057
KEYWORDS   3 of 27
SEGMENT    Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 3058)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE      Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE    20283944
PUBMED     10811918
REFERENCE  2 (bases 1 to 3058)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE      Direct Submission
JOURNAL    Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES   Location/Qualifiers
source     1..3058
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20p13"
            810..923
            /gene="ATRN"
            /number=3
exon       495 GCCTAATAGATATAGAGACTCGTTTCATCATTCATTGCTACAGAGTGTAGTGGACCA 554
            |||
            810 GCCTAATAGATATAGAGACTCGTTTCATCATTCATTGCTACAGAGTGTAGTGGACCA 869
            |||
            555 TTATATGTTTATGATGGGACTCAATTATGACCGCTAGTCTGCATTAG 608
            |||
            870 TTATATGTTTATGATGGGACTCAATTATGACCGCTAGTCTGCATTAG 923
            |||
RESULT 46
AF273610
LOCUS      AF273610      237 bp      mRNA      linear      MAM 09-JUL-2000
DEFINITION Sus scrofa mahogany mRNA, partial cds.
ACCESSION AF273610
VERSION    AF273610.1 GI:8980838
KEYWORDS   2 (bases 1 to 237)
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
REFERENCE  1 (bases 1 to 237)
AUTHORS   Matteri, R.L. and Dyer, C.J.
TITLE      Sus scrofa mahogany mRNA
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 237)
AUTHORS   Matteri, R.L. and Dyer, C.J.
TITLE      Direct Submission
JOURNAL    Submitted (01-JUN-2000) Animal Physiology Research Unit, USDA -
Agricultural Research Service, University of Missouri, 920 East
Campus Dr., Room S-143 ASRC, Columbia, MO 65211, USA
FEATURES   Location/Qualifiers
source     1..1592
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20p13"
            747..854
            /gene="ATRN"
            /number=28
exon       2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 2304
            |||
            125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 184
            |||
            2305 CAGAACTGCAGTGGGAGCCCGGATCAGGATCAGGATGCGCTGCCGAAA 2357
            |||
            185 CAGAACTGCAGTGGGAGCCCGGATCAGGATGCGCTGCCGAAA 237
            |||
Query Match 2.6%; Score 113; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.6e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 2304
Db 125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 184
QY 2305 CAGAACTGCAGTGGGAGCCCGGATCAGGATCAGGATGCGCTGCCGAAA 2357
Db 185 CAGAACTGCAGTGGGAGCCCGGATCAGGATGCGCTGCCGAAA 237
RESULT 47
H009384S25
LOCUS      H009384S25      1592 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 28.
ACCESSION AF218913
VERSION    AF218913.1 GI:8118079
KEYWORDS   25 of 27
SEGMENT    Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1592)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE      Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE    20283944
PUBMED     10811918
REFERENCE  2 (bases 1 to 1592)
AUTHORS   Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE      Direct Submission
JOURNAL    Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES   Location/Qualifiers
source     1..1592
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20p13"
            747..854
            /gene="ATRN"
            /number=28
exon       2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 2304
            |||
            125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCTGCAGGAGCTGTGCTCTGGAC 184
            |||
            2305 CAGAACTGCAGTGGGAGCCCGGATCAGGATCAGGATGCGCTGCCGAAA 2357
            |||
            185 CAGAACTGCAGTGGGAGCCCGGATCAGGATGCGCTGCCGAAA 237
            |||
Query Match 2.6%; Score 110; DB 9; Length 1592;
Best Local Similarity 100.0%; Pred. No. 5.4e-47;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3941 AGCAACTTCTTCGAGATGCAACAGATGGCCAGCGTCCCTTGGCTCTGTAATGTCG 4000
Db 745 AGCAACTTCTTCGAGATGCAACAGATGGCCAGCGTCCCTTGGCTCTGTAATGTCG 804
QY 4001 CTTTGAACAGATGAGAGCTCTGATCTATTGGGGGAGTATAAG 4050
Db 805 CTTTGAACAGATGAGAGCTCTGATCTATTGGGGGAGTATAAG 854

RESULT 48
H009384S17 1675 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 20.
DEFINITION AF218905
ACCESSION AF218905.1 GI:8118071
KEYWORDS
SEGMENT 17 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1675)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918

REFERENCE 2 (bases 1 to 1675)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES
source
1..1675
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
1343..1433
/gene="ATRN"
/number=7

exon
1173 TAGATATGTCATCTTTGGCATATACAGG 1204
1403 TAGATATGTCATCTTTGGCATATACAGG 1434

ORIGIN
Query Match 2.2%; Score 94; DB 9; Length 1675;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3323 CATGCAAGTGCATGGGCAACGGCTCTGTGCAACACCAACGCGGCAAGTCTTGTGCA 3382
Db 833 CATGCAAGTGCATGGGCAACGGCTCTGTGCAACACCAACGCGGCAAGTCTTGTGCA 892

QY 3383 CCACCAAGGGCGTCAAGGGGAGCAGTGCCAGCT 3416
Db 893 CCACCAAGGGCGTCAAGGGGAGCAGTGCCAGCT 926

RESULT 49
H009384S07 1575 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 7.
DEFINITION AF218895
ACCESSION AF218895.1 GI:8118061
KEYWORDS
SEGMENT 7 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1575)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918

REFERENCE 2 (bases 1 to 1575)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES
source
1..1575
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
833..926
/gene="ATRN"
/number=20

exon
1113 GATGACCTTGTCTTAGGGAGTGGCTTCCATTAACCGTCTGTGAACAATGGTGT 1172
1343 GATGACCTTGTCTTAGGGAGTGGCTTCCATTAACCGTCTGTGAACAATGGTGT 1402

ORIGIN
Query Match 2.1%; Score 92; DB 9; Length 1575;
Best Local Similarity 100.0%; Pred. No. 3e-37;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1113 GATGACCTTGTCTTAGGGAGTGGCTTCCATTAACCGTCTGTGAACAATGGTGT 1172
Db 1343 GATGACCTTGTCTTAGGGAGTGGCTTCCATTAACCGTCTGTGAACAATGGTGT 1402

QY 1173 TAGATATGTCATCTTTGGCATATACAGG 1204
Db 1403 TAGATATGTCATCTTTGGCATATACAGG 1434

RESULT 50
H009384S02 3711 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 2.
DEFINITION AF218890
ACCESSION AF218890.1 GI:8118056
KEYWORDS
SEGMENT 2 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3711)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918

REFERENCE 2 (bases 1 to 3711)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES
source
1..3711
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"

TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

FEATURES
source
1..1704 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
873..951
/gene="ATRN"
/number=23

exon

ORIGIN
Query Match 1.8%; Score 79; DB 9; Length 1704;
Best Local Similarity 100.0%; Pred.No. 3.2e-30;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3565 CAAAACAGGATTGGACATGTTTCATCATGCTCCAGAAATTTCAACCTCAACATCACC 3624
|||||
DB 873 CAAAACAGGATTGGACATGTTTCATCATGCTCCAGAAATTTCAACCTCAACATCACC 932
|||||

QY 3625 TGGGCTGCCAGTTTCTCAG 3643
|||||
DB 933 TGGGCTGCCAGTTTCTCAG 951
|||||

RESULT 54
LOCUS H009384S12
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 13.
ACCESSION AF218900
VERSION AF218900.1 GI:8118066
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1267)
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE
AUTHORS 2 (bases 1 to 1267)
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

FEATURES
source
1..1267 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
728..849
/gene="ATRN"
/number=13

exon

ORIGIN
Query Match 1.7%; Score 75; DB 9; Length 1267;
Best Local Similarity 100.0%; Pred.No. 4.9e-28;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2140 ACAGCCACCAATGACTGGTGGTCAATGACCATTTGTCTCCAGGACCAACGAGC 2199
|||||

Db 775 ACAGCCACCAATGACTGGTGGTCAATGACCATTTGTCTCCAGGACCAACGAGC 834
|||||

QY 2200 TGCTCAGAGGCCAG 2214
|||||

Db 835 TGCTCAGAGGCCAG 849
|||||

RESULT 55
LOCUS H009384S23
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 26.
ACCESSION AF218911
VERSION AF218911.1 GI:8118077
KEYWORDS
SEGMENT
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2214)
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE
AUTHORS 2 (bases 1 to 2214)
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

FEATURES
source
1..2214 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
1348..1409
/gene="ATRN"
/number=26

exon

ORIGIN
Query Match 1.5%; Score 65; DB 9; Length 2214;
Best Local Similarity 100.0%; Pred.No. 1.2e-22;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3799 CAGATTGCCTTCTCAGCAGCAGCAATTTATGGACCTGGTACAGTTCTTGTGACTTTC 3858
|||||

Db 1345 CAGATTGCCTTCTCAGCAGCAGCAATTTATGGACCTGGTACAGTTCTTGTGACTTTC 1404
|||||

QY 3859 TTCAG 3863
|||||

Db 1405 TTCAG 1409
|||||

RESULT 56
LOCUS AB062913
DEFINITION Mesocricetus auratus Atrn mRNA for attractin, complete cds.
ACCESSION AB062913
VERSION AB062913.1 GI:16930100
KEYWORDS
SOURCE Mesocricetus auratus (golden hamster)
ORGANISM Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.

REFERENCE 1

AUTHORS Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and Ushijima,T.
TITLE Insertional mutation of the Attractin gene in the black tremor hamster
JOURNAL Mamm Genome 13 (1), 36-40 (2002)
MEDLINE 21635551
PUBMED 11773967
REFERENCE 2 (bases 1 to 4284)
AUTHORS Kuramoto,T. and Ushijima,T.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp, URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511, Fax:81-3-5565-1753)
FEATURES
source Location/Qualifiers
1..4284 /organism="Mesocricetus auratus"
/mol_type="mRNA"
/db_xref="taxon:10036"
1..4284 /gene="Atrn"
1..4284 /gene="Atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB72012.1"
/db_xref="GI:16930101"
/translation="MVAAAATALLRGSTAAAPAGRKVEQYRLCAATGAWRPGPR
AGLCLPWSRALPPPLLLLLSLLPLPREAABAAVAAVSGSAAAEAKCEDRP
CVNGRNPCTGCGVCPAGVGCOCQGRFRLTGGSGFVTDGPNKYKTKTWLI
EGOPNRLMFLRNFHFAECSDHLYVVDGDSIVAPLIAAFGLIIPERDGNETAPVT
VTSGYALLHFFSDAAYNLGTNITYNFDMPNCSGRGCKSSNSNTVECESENK
GESCDPHCTDNCGFPHRGICNASDTRGCSFHWQPGCSIIPVPAQNSFWTKEEYSD
FLKPRASHKAVNGNIMVVGYNFNSDYSMLYDLASREPLNHSVNVVVRIG
HSLAHKDKIYMGKIDSTGNVFNELRVFHIHNSWLLTPKAEQYAVGSHAHV
TLATGRVLMVIFGHCPLYGYSIVQYEDLEKNTWILTKDSRPFRYLHTAVISGTMV
TKALYHVGKYKAFSANKYRLADDLRYVDVDTQMTILKDSRFRYLHTAVISGTMV
EGNTNDTSMHGAKCFSSDFMAYDIADCRWSVLRPELHDVNRHSAVLNYS
YVFGGNSLLSVDLVFTSEQDAHREAAACVAGFIRCLWDVTSRCSWELATEE
QAEKLKSECFSKRTLDRCDQHTDCYCTANTNDCHNDCVFNHSCSTEGQIS
KYDNCPCPNMYCNKRTSCRSALDQNCQWEPNQCIALPENICGNHVLGNSCL
KITAKENYDNALSKRNHAFSLTSQKVELVLRKQLRMOSQSMKLTLPWG
LRKINYSWCWEDMSPNTSLLOWMPSRSDAGFCGILSEPTSGRLKAATCINPLNG
VCEPPANSAKQCTPCALRTAGCTSSSSECCWCSNMKQCVDSNAYVASFPFGOC
EWYTMSSCPFPCSGYCTCHLCLEQPCGCTDPSNTGKCKIEGSKYKGPMPQSA
TGNYVPOPLNNSMCLSDSYNWSFHCPCQCNHSGKINQISCKEKNLTGKHCE
TGISGFYDPTNGKQCPCKNGHSLCNTNTGKCTTGVKGDCEQCEVNRVQG
NPLKTCYTYLLIDYQTFSLQSDDRYTAINFATPDQNRDNFDMFINASKFNLI
ITWATSPAGTQGEVVPVSKNIKEYKDSFNKEKDFRNHNTFFVYVSNFTWPI
KIOTARSOHNFMDLVQFVTFPFCFLSLLVAVVKIKOSCAWRREQLIREMOQ
MASRPPASVNALETDEEPDLIGGSIKTPKLEPCFGKKAALVSVFRLPRGL
GIPPGQSGLAVALVDISQMPVIVYKESGAVNRKQOPPAQPGTCTI"

AUTHORS AF119821, 1 GI:4585306
TITLE Mus musculus (house mouse)
JOURNAL Mus musculus
MEDLINE 10086336
PUBMED 99844160
REFERENCE 2 (bases 1 to 4313)
AUTHORS Schloeman,S.P., Duke-Cohan,J.S. and Barsh,G.S.
TITLE The mouse mahogany locus encodes a transmembrane form of human attractin
JOURNAL Nature 398 (6723), 152-156 (1999)
MEDLINE 99844160
PUBMED 10086336
REFERENCE 2 (bases 1 to 4313)
AUTHORS Schloeman,S.P., Duke-Cohan,J.S. and Barsh,G.S.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1999) HMI, Stanford, Beckman Center B271A, Stanford, CA 94305-5323, USA
FEATURES
source Location/Qualifiers
1..4313 /organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="2"
1..4313 /gene="Mgca"
/note="mahogany candidate"
27..4313 /gene="Mgca"
/note="mahogany"
/codon_start=1
/product="attractin"
/protein_id="AAD25372.1"
/db_xref="GI:4585307"
/translation="MVAVAAATEARLGSTTATAAPAGRKGRHREPTATGAWRPG
PRARCLPRVLRALPPPLPLPLSLLPLPREAABAAVAAVSGSAAAEAKCEDR
PCVNGRNPCTGCGVCPAGVGCOCQGRFRLTGGSGFVTDGPNKYKTKTWLI
EGEPNVLRLRNFHFAECSDHLYVVDGDSIVAPLIAAFGLIIPERDGNETAPVT
VTSGYALLHFFSDAAYNLGTNITYNFDMPNCSGRGCKSSNSNTVECESENK
GESCDPHCTDNCGFPHRGICNASDTRGCSFHWQPGCSIIPVPAQNSFWTKEEYSD
FLKPRASHKAVNGNIMVVGYNFNSDYSMLYDLASREPLNHSVNVVVRIG
HSLAHKDKIYMGKIDSTGNVFNELRVFHIHNSWLLTPKAEQYAVGSHAHV
TLATGRVLMVIFGHCPLYGYSIVQYEDLEKNTWILTKDSRFRYLHTAVISGTMV
TKALYHVGKYKAFSANKYRLADDLRYVDVDTQMTILKDSRFRYLHTAVISGTMV
EGNTNDTSMHGAKCFSSDFMAYDIADCRWSVLRPELHDVNRHSAVLNYS
YVFGGNSLLSVDLVFTSEQDAHREAAACVAGFIRCLWDVTSRCSWELATEE
QAEKLKSECFSKRTLDRCDQHTDCYCTANTNDCHNDCVFNHSCSTEGQIS
KYDNCPCPNMYCNKRTSCRSALDQNCQWEPNQCIALPENICGNHVLGNSC
KITAKENYDNALSKRNHAFSLTSQKVELVLRKQLRMOSQSMKLTLPW
GLRKNYSWCWEDMSPNTSLLOWMPSRSDAGFCGILSEPTSGRLKAATCINPLNG
VCEPPANSAKQCTPCALRTAGCTSSSSECCWCSNMKQCVDSNAYVASFPFGOC
EWYTMSSCPFPCSGYCTCHLCLEQPCGCTDPSNTGKCKIEGSKYKGPMPQSA
TGNYVPOPLNNSMCLSDSYNWSFHCPCQCNHSGKINQISCKEKNLTGKHCE
TGISGFYDPTNGKQCPCKNGHSLCNTNTGKCTTGVKGDCEQCEVNRVQG
NPLKTCYTYLLIDYQTFSLQSDDRYTAINFATPDQNRDNFDMFINASKFNLI
ITWATSPAGTQGEVVPVSKNIKEYKDSFNKEKDFRNHNTFFVYVSNFTWPI
KIOTARSOHNFMDLVQFVTFPFCFLSLLVAVVKIKOSCAWRREQLIREMOQ
MASRPPASVNALETDEEPDLIGGSIKTPKLEPCFGKKAALVSVFRLPRGL
GIPPGQSGLAVALVDISQMPVIVYKESGAVNRKQOPPAQPGTCTI"

Query Match 1.5%; Score 64; DB 10; Length 4284;
Best Local Similarity 100.0%; Pred. No. 3.8e-22;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1607 GATATGATGGATACCCAGATGGACCATCTTAAAGGACGCCGATTTTCGGTTACT 1666
DB 1601 GATATGATGGATACCCAGATGGACCATCTTAAAGGACGCCGATTTTCGGTTACT 1660
QY 1667 TGCA 1670
DB 1661 TGCA 1664
RESULT 57
AF119821
LOCUS AF119821 4313 bp mRNA linear ROD 16-APR-1999
DEFINITION Mus musculus attractin (Mgca) mRNA; complete cds.
ACCESSION AF119821

AF119821.1 GI:4585306
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4313)
Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
Schloeman,S.P., Duke-Cohan,J.S. and Barsh,G.S.
The mouse mahogany locus encodes a transmembrane form of human
attractin
Nature 398 (6723), 152-156 (1999)
99844160
10086336
2 (bases 1 to 4313)
Schloeman,S.P., Duke-Cohan,J.S. and Barsh,G.S.
Direct Submission
Submitted (12-JAN-1999) HMI, Stanford, Beckman Center B271A,
Stanford, CA 94305-5323, USA
Location/Qualifiers
1..4313 /organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="2"
1..4313 /gene="Mgca"
/note="mahogany candidate"
27..4313 /gene="Mgca"
/note="mahogany"
/codon_start=1
/product="attractin"
/protein_id="AAD25372.1"
/db_xref="GI:4585307"
/translation="MVAVAAATEARLGSTTATAAPAGRKGRHREPTATGAWRPG
PRARCLPRVLRALPPPLPLPLSLLPLPREAABAAVAAVSGSAAAEAKCEDR
PCVNGRNPCTGCGVCPAGVGCOCQGRFRLTGGSGFVTDGPNKYKTKTWLI
EGEPNVLRLRNFHFAECSDHLYVVDGDSIVAPLIAAFGLIIPERDGNETAPVT
VTSGYALLHFFSDAAYNLGTNITYNFDMPNCSGRGCKSSNSNTVECESENK
GESCDPHCTDNCGFPHRGICNASDTRGCSFHWQPGCSIIPVPAQNSFWTKEEYSD
FLKPRASHKAVNGNIMVVGYNFNSDYSMLYDLASREPLNHSVNVVVRIG
HSLAHKDKIYMGKIDSTGNVFNELRVFHIHNSWLLTPKAEQYAVGSHAHV
TLATGRVLMVIFGHCPLYGYSIVQYEDLEKNTWILTKDSRFRYLHTAVISGTMV
TKALYHVGKYKAFSANKYRLADDLRYVDVDTQMTILKDSRFRYLHTAVISGTMV
EGNTNDTSMHGAKCFSSDFMAYDIADCRWSVLRPELHDVNRHSAVLNYS
YVFGGNSLLSVDLVFTSEQDAHREAAACVAGFIRCLWDVTSRCSWELATEE
QAEKLKSECFSKRTLDRCDQHTDCYCTANTNDCHNDCVFNHSCSTEGQIS
KYDNCPCPNMYCNKRTSCRSALDQNCQWEPNQCIALPENICGNHVLGNSC
KITAKENYDNALSKRNHAFSLTSQKVELVLRKQLRMOSQSMKLTLPW
GLRKNYSWCWEDMSPNTSLLOWMPSRSDAGFCGILSEPTSGRLKAATCINPLNG
VCEPPANSAKQCTPCALRTAGCTSSSSECCWCSNMKQCVDSNAYVASFPFGOC
EWYTMSSCPFPCSGYCTCHLCLEQPCGCTDPSNTGKCKIEGSKYKGPMPQSA
TGNYVPOPLNNSMCLSDSYNWSFHCPCQCNHSGKINQISCKEKNLTGKHCE
TGISGFYDPTNGKQCPCKNGHSLCNTNTGKCTTGVKGDCEQCEVNRVQG
NPLKTCYTYLLIDYQTFSLQSDDRYTAINFATPDQNRDNFDMFINASKFNLI
ITWATSPAGTQGEVVPVSKNIKEYKDSFNKEKDFRNHNTFFVYVSNFTWPI
KIOTARSOHNFMDLVQFVTFPFCFLSLLVAVVKIKOSCAWRREQLIREMOQ
MASRPPASVNALETDEEPDLIGGSIKTPKLEPCFGKKAALVSVFRLPRGL
GIPPGQSGLAVALVDISQMPVIVYKESGAVNRKQOPPAQPGTCTI"

Query Match 1.4%; Score 62; DB 10; Length 4313;
Best Local Similarity 100.0%; Pred. No. 4.6e-21;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3130 TCACGATGTGTAGAGACAGACAGATACAACTGGTCTTTTCATTTCACATGTCACCTTGC 3189
DB 3153 TCACGATGTGTAGAGACAGACAGATACAACTGGTCTTTTCATTTCACATGTCACCTTGC 3212
QY 3190 CA 3191

[illegible]

CNGHSLCNTNTKCTCTCTGKVGKDECOLCEVENRYQGNPLKGTCTVYTLIDYQFTFS
LSQEDRYVTAINFATPDQONEDLDVFINASKNENLNTWATSPAGTGTDEVEPV
SKTIEKQDSFNEKFDNRHNPIFFVYSNFTWPKIKIQAPOSQSNPMDLVQFV
TFSCFUSLLVAAVVKIKOSCWAGRREQLLEQWQASRFASFVVALETDSEPP
DLIGSFKITVPKPIALEPCFNGKAAVLSVFLRPLRGGLGIPPGQSLAVASALVDIS
QCMPIVYKESGAVNRKQPPPAQPGTCT

ORIGIN

Query Match 1.4%; Score 62; DB 10; Length 5683;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3130 TCCAGCATGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGC 3189
Db 2021 TCCAGCATGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGC 2080
Qy 3190 CA 3191
Db 2081 CA 2082

RESULT 60
AF116897 AF116897 6083 bp mRNA linear ROD 20-MAR-1999
LOCUS Mus musculus mahogany protein mRNA, complete cds.
DEFINITION AF116897
ACCESSION AF116897
VERSION AF116897.1 GI:4454560
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 6083)
Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
DiRocco,L., Holmgren,L., Montagnon,J., Bork,P., Huszar,D.,
Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
The mahogany protein is a receptor involved in suppression of
obesity

JOURNAL Nature 398 (6723), 148-152 (1999)
MEDLINE 99184159
PUBMED 10086355
REFERENCE 2 (bases 1 to 6083)
Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
DiRocco,L., Holmgren,L., Montagnon,J., Bork,P., Huszar,D.,
Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
Direct Submission
Submitted (28-DEC-1998) Sequence Analysis, Millennium
Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA

FEATURES
source
Location/Qualifiers
1..6083
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
70..4488
/note="MG"

CDS
1..codon_start=1
/product="mahogany protein"
/protein_id="AA020947.1"
/db_xref="GI:4454561"
/translation="MVAVAAAAAATEARLGRSTTTTAAAGKRGQRHRCPTATGAWRPG
PRALCPLRVLSRALPPPLPLPFLSLLPLPREAATAAATAVSGSAAAEKCDR
PCVNGRCNPGTCQVCTGPTGQCCOHCGRFELTSGSGFVDTGPGNFKYKTKTWL
IEGQNRIMRLNFHFAFECSDWHLIVYDGSIYAPLIAAFSGLIVPERGNTAPEV
TVTSGYALLHFFSDAAYNLNLTGNITYNFDMPCFNCSGRGCKSSNVAECCSENW
KGESCDIPHCNDGCFPHRGICNASDTGCGSPFHQPGCGSIPVPAQFWRREYS
DLKLPRAHKAIVNGNIMVVGYNFNSHDYSMLAYDLTSREWLPLNHNVSVVRY
GSLALHKKLYMGKIDSTGNVTNELRVFHINESWLLTPAKQDQAVVGSAAH
VTLASGRVWLVIHQHCPYLCYISVQVQYDLKNTWLSILHTQGLVQGYGSHVYDD
RTKALVHGKTKASAKYRLADLIRYDVDTQMTWTLKDSRFRYIHTAVISGTM
VFGNTHNDTSMHGAKCFSSDFMAYDIACDRVSLVPRPELHDVNRFGHSAVLNST
MYVFGGNSLLSDVLTFTSQCDHARSEAAVAAAGPGIRCLMDTQSSRCTSMELATE

BOAEKLSCEFSKRTLDHRCDOHQDTDCYSTANTNDCHWCHDCHVPVNHSTEQOISI
AKYESCPKDNMYCYNKKTSRSCALDONCOWEPENQECIALPENICGWHLVGNSC
LKITAKENYDNAKLSCHNNAFLASLTQKKVEFLKQLRLMOSSQMSKLTLPWV
GLRLNVSCHWEDMSPNTSILLOMPSEBPDAGFCGLSBSPTGKAACINPLNG
SVCEIPANHSKQCKIPCALRTACGCTSSSCWCSNMKQCVDSNAYVASFPPGQC
MEWYTMSPENCPSGYCTCSHLEGPCCGCTDPSNTGKCKLEGSKVGPVKMPSQA
SAGNYVQPLLNSMCLSDSYNWSFHCPACQCNHSHKCIQSICEKEDLTGKHHC
ETCISGYDPTNGKQCPCKCNHSHNTNTGKCFCTTGKVGDECOLCEVENRYQ
GNPLKGTCTVYTLIDYQFTFSLOEDDRYTAINFATPDQONEDLDVFINASKNENL
NITWATSPAGTGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGC 3189
IKIQAPOSQSNPMDLVQFVTFSCFUSLLVAAVVKIKOSCWAGRREQLLEQWQASRFASFVVALETDSEPP
QCMPIVYKESGAVNRKQPPPAQPGTCT
LTVGWSKLTSTGEAAGRQAWGRRLGLEASNSCA"

ORIGIN

Query Match 1.4%; Score 62; DB 10; Length 6083;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3130 TCCAGCATGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGC 3189
Db 3196 TCCAGCATGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTCTCCAGCTTGC 3255
Qy 3190 CA 3191
Db 3256 CA 3257

RESULT 61

AB038387 AB038387 8739 bp mRNA linear ROD 18-JAN-2001
LOCUS Rattus norvegicus atrn mRNA for attractin, complete cds.
DEFINITION AB038387
ACCESSION AB038387
VERSION AB038387.1 GI:12275307
KEYWORDS attractin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (sites)
AUTHORS Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
and Serikawa,T.
TITLE Attractin/mahogany/zitter plays a critical role in myelination of
the central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
MEDLINE 21143347
PUBMED 11209055

REFERENCE 2 (bases 1 to 8739)
AUTHORS Kuramoto,T., Serikawa,T. and Ushijima,T.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division, Tsukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp,
Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)

FEATURES

source
Location/Qualifiers
1..8739
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="WTC"
/db_xref="taxon:10116"
1..8739
/gene="atrn"
61..4359
/gene="atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB21017.1"
/db_xref="GI:12275308"
/translation="MVAAAAAATEARLGRYTTATAAGWKERQHRPCATGAWRP
WPAAGCLPRVLSRALSPPLPLPFLSLLPLPREAATAAATAVSGSAAAEK


```

REFERENCE 1 (bases 1 to 966)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 966)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
COMMENT On Nov 24, 2000 this sequence version replaced gi:8118072.
FEATURES
source
Location/Qualifiers
1..966
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
49..101
/gene="ATRN"
/number=21
exon
Query Match 1.2%; Score 53; DB 9; Length 966;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3417 ATGTGAGTGAATAATCATACCAAGGAACCCCTCTCAGGACATGTTATT 3469
|||||
DB 49 ATGTGAGTGAATAATCATACCAAGGAACCCCTCTCAGGACATGTTATT 101
|||||

RESULT 65
AR164814
LOCUS AR164814 1051 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 10 from patent US 6274339.
ACCESSION AR164814
VERSION AR164814.1 GI:16238023
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1051)
AUTHORS Moore, K. and Nagle, D. Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 10 14-AUG-2001;
FEATURES
source
Location/Qualifiers
1..1051
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 1.2%; Score 53; DB 6; Length 1051;
Best Local Similarity 100.0%; Pred. No. 4e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 367 GGCTGGTGGGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAATCGG 419
|||||
DB 571 GGCTGGTGGGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAATCGG 623
|||||

RESULT 66
AB062914S3
LOCUS AB062914S3 1463 bp DNA linear ROD 16-JAN-2002
DEFINITION Mesocricetus auratus Atrn gene for attractin, exon 25, 26.
ACCESSION AB062916
VERSION AB062916.1 GI:16930138

```

```

KEYWORDS 3 of 5
SEGMENT Mesocricetus auratus (golden hamster)
SOURCE Mesocricetus auratus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
REFERENCE 1
AUTHORS Kuramoto, T., Nomoto, T., Fujiwara, A., Mizutani, M., Sugimura, T. and
Ushijima, T.
TITLE Insertional mutation of the Attractin gene in the black tremor
hamster
JOURNAL Mamm. Genome 13 (1), 36-40 (2002)
MEDLINE 21635551
PUBMED 11773967
REFERENCE 2 (bases 1 to 1463)
AUTHORS Kuramoto, T. and Ushijima, T.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
Chuo-Ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp,
URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
Fax:81-3-5565-1753)
FEATURES
source
Location/Qualifiers
1..1463
/organism="Mesocricetus auratus"
/mol_type="genomic DNA"
/db_xref="taxon:10036"
426..487
/gene="Atrn"
/product="attractin"
/note="CDS is reported in Acc# AB062918"
/number=25
repeat_region 604..663
/rpt_type=tandem
/rpt_unit="ga"
766..844
/gene="Atrn"
/product="attractin"
/note="CDS is reported in Acc# AB062918"
/number=26
exon
ORIGIN
Query Match 1.2%; Score 50; DB 10; Length 1463;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3814 CACCACAGCAATTTATGGACCTGGTACAGTTCTTCGACTTCTTCAG 3863
|||||
DB 438 CACCACAGCAATTTATGGACCTGGTACAGTTCTTCGACTTCTTCAG 487
|||||

RESULT 67
H009384S19
LOCUS H009384S19 1781 bp DNA linear PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 22.
ACCESSION AF218907
VERSION AF218907.1 GI:8118073
KEYWORDS
SEGMENT 19 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1781)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1781)

```

AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
 TITLE Direct Submission
 JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES
 source Location/Qualifiers
 1..1781
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20p13"
 802..896
 /gene="ATR"n
 /number=22

exon

ORIGIN
 Query Match 1.2%; Score 50; DB 9; Length 1781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-14;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3515 AAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTGACGAA 3564
 |||||
 Db 847 AAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTGACGAA 896
 |||||

RESULT 68
 AF301008
 LOCUS Bos taurus attractin mRNA, partial cds. MAM 26-SEP-2000
 DEFINITION
 ACCESSION AF301008
 VERSION AF301008.1 GI:10304438
 KEYWORDS
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 315)
 AUTHORS Glad, T.W., Granholm, N.H., Westby, C.A., Marshall, D.M. and Kraft, E.
 TITLE Bovine mahogany/attractin cDNA, including exons 26-29
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 315)
 AUTHORS Glad, T.W., Granholm, N.H., Westby, C.A., Marshall, D.M. and Kraft, E.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Bio/Micro, South Dakota State University, 2140D, 252 NPB, Brookings, SD 57007, USA

FEATURES
 source Location/Qualifiers
 1..315
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue type="cerebrum"
 <1..>315
 /note="mahogany"
 /codon_start=2
 /product="attractin"
 /protein_id="AAG16254.1"
 /db_xref="GI:10304439"
 /translation="HSNFMDLVQFPVTFVFFSCFLSLLLVAAVWVKIKSCWASRRREOL LREQMARSFPASVNVVALETDEBPDLIGSSIKVTPKPIALEPCFNKAVLVSVFR LPR"

CDS

ORIGIN
 Query Match 1.1%; Score 49; DB 4; Length 315;
 Best Local Similarity 100.0%; Pred. No. 6.8e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3878 TGCTCCTGGTGGCTGCTGCTGCTGGTGGAGATCAACAAGTTGTTGGC 3926
 |||||
 Db 63 TGCCTCCTGGTGGCTGCTGCTGCTGGTGGAGATCAACAAGTTGTTGGC 111
 |||||

AUTHORS Kuramoto, T., Nomoto, T., Fujiwara, A., Mizutani, M., Sugimura, T. and Ushijima, T.
 TITLE Insertional mutation of the Attractin gene in the black tremor hamster
 JOURNAL Mamm. Genome 13 (1), 36-40 (2002)
 MEDLINE 21635551
 PUBMED 11773967
 REFERENCE 2 (bases 1 to 4862)
 AUTHORS Kuramoto, T. and Ushijima, T.
 TITLE Direct Submission
 JOURNAL Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp, URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511, Fax:81-3-5565-1753)

FEATURES
 source Location/Qualifiers
 1..4862
 /organism="Mesocricetus auratus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10036"
 594..672
 /gene="Atrn"
 /product="attractin"
 /note="CDS is reported in Acc# AB062918"
 /number=23
 2976..3133
 /gene="Atrn"
 /product="attractin"
 /note="CDS is reported in Acc# AB062918"
 /number=24
 3782..3886
 /gene="Atrn"
 /rpt_type=tandem
 3782..3886
 /rpt_unit="aag"

ORIGIN
 Query Match 1.1%; Score 48; DB 10; Length 4862;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3754 ACTTCTTTGTTTATGTCAGTAATTCACCTGGCCCAAAATTCAG 3801
 |||||
 Db 3086 ACTTCTTTGTTTATGTCAGTAATTCACCTGGCCCAAAATTCAG 3133
 |||||

RESULT 70
 E26747
 LOCUS Structure and function of novel biorhythm marker gene.
 DEFINITION
 ACCESSION E26747
 VERSION E26747.1 GI:13026325
 KEYWORDS JP 1999169185-A/1.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1012)
AUTHORS Mario,I., Yutaka,S., Satoru,S. and Mikio,K.
TITLE Structure and function of novel bionhythm marker gene
JOURNAL Patent: JP 1999169185-A 1 29-JUN-1999;
AGENCY OF IND SCIENCE & TECHNOL,OTSUKA PHARMACEUT CO LTD
COMMENT OS Rattus sp. (Wistar rat)
PN JP 1999169185-A/1
PD 29-JUN-1999
PF 12-DEC-1997 JP 1997362890
PR MARIO ISHIDA, YUTAKA SADAKANE, SATORU SUZUKI, MIKIO KIKUCHI PC
C12N15/09,C07K14/47,C12Q1/68/C12N5/10,C12P21/02,C12N15/09, PC
C12R1.91),
PC (C12N5/10,C12R1.91), (C12P21/02,C12R1.91),C12N15/00,C12N5/00,
PC (C12N15/00,C12R1.91), (C12N5/00,C12R1.91)
CC Strandedness: Single;
CC Topology: Linear;
FH Key 1. .720.
FT CDS Location/Qualifiers
FEATURES
source
1..1012
/organism="Rattus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10118"
ORIGIN
Query Match 1.1%; Score 47; DB 6; Length 1012;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2698 AGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGG 2744
DB 466 AGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGG 512
RESULT 71
AB06968S15
LOCUS Rattus norvegicus (Norway rat)
DEFINITION Rattus norvegicus Atrn gene for membrane attractin, exon 16.
ACCESSION AB049236
VERSION AB049236.1 GI:12275377
KEYWORDS
SEGMENT
SOURCE
ORGANISM
REFERENCE
AUTHORS Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
and Serikawa,T.
TITLE Attractin/mahogany/zitter plays a critical role in myelination of
the central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
MEDLINE 21143347
PUBMED 11209055
REFERENCE 2 (bases 1 to 1207)
AUTHORS Kuramoto,T. and Ushijima,T.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division; Teukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp,
Tel:81-33542-2511, Fax:81-5565-1753)
LOCATION/Qualifiers
FEATURES
source
1..1207
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="3"
488..707
/gene="Atrn"
/product="membrane attractin"
exon

ORIGIN
Query Match 1.1%; Score 47; DB 10; Length 1207;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2698 AGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGG 2744
DB 641 AGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGG 687
RESULT 72
AR164813
LOCUS Sequence 8 from patent US 6274339.
DEFINITION AR164813
ACCESSION AR164813
VERSION AR164813.1 GI:16238022
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2419)
AUTHORS Moore,K. and Magle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 8 14-AUG-2001;
FEATURES
source
1..2419
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 1.1%; Score 47; DB 6; Length 2419;
Best Local Similarity 100.0%; Pred. No. 6.5e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1624 CAGATGTGACCACTTCTTAAGGACAGCCGATTTTCGGTACTTGA 1670
DB 1890 CAGATGTGACCACTTCTTAAGGACAGCCGATTTTCGGTACTTGA 1936
RESULT 73
AL954766.2
WPCOMMENT
Sequence split into 4 fragments LOCUS AL954766 Accession AL954766
Fragment Name Begin End
AL954766_1 1 110000
AL954766_2 100001 210000
AL954766_3 200001 310000
AL954766_4 300001 356753
Continuation (3 of 4) of AL954766 from base 200001 (AL954766 Mus musculus chromosome X
Query Match 1.1%; Score 47; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 4.1e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3829 ATGGACCTGTGTACAGTTCTTCGTGACTTCTTCAGTTGTTTCCTCTC 3875
DB 60794 ATGGACCTGTGTACAGTTCTTCGTGACTTCTTCAGTTGTTTCCTCTC 60840
RESULT 74
BX004762.0/c
WPCOMMENT
Sequence split into 5 fragments LOCUS BX004762 Accession BX004762
Fragment Name Begin End
BX004762_0 1 110000
BX004762_1 100001 210000
BX004762_2 200001 310000
BX004762_3 300001 410000
BX004762_4 400001 442165

LOCUS BX004762 442165 bp DNA linear HTG 12-APR-2003
 DEFINITION Mus musculus chromosome X clone RP23-146C4, WORKING DRAFT SEQUENCE,
 4 unordered pieces.
 ACCESSION BX004762
 VERSION HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 442165)
 Chapman, J.
 Direct Submission
 Submitted (11-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Apr 13, 2003 this sequence version replaced gi:29125168.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: BM146C4
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 229053 bases at least Q40
 Consensus quality: 229393 bases at least Q30
 Consensus quality: 229535 bases at least Q20
 Insert size: 441865; sum-of-contigs
 Insert size: 188914; 6.6% error; agarose-fp
 Quality coverage: 3.34x in Q20 bases; sum-of-contigs Quality
 coverage: 7.96x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 48919: contig of 48919 bp in length
 * 48920 49019: gap of 100 bp
 * 49020 130781: contig of 81762 bp in length
 * 130782 130881: gap of 100 bp
 * 130882 229564: contig of 95083 bp in length
 * 229565 230064: gap of 100 bp
 * 230065 442165: contig of 21201 bp in length.
 FEATURES
 source
 1..442165
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="X"
 /clone="RP23-146C4"
 /clone_lib="RPCI-23"
 1..48919
 /note="assembly fragment:00147"
 fragment_chain:1
 49020..130781
 /note="assembly fragment:00500"
 fragment_chain:1
 130882..229564
 /note="assembly fragment:00870"
 fragment_chain:1
 230065..442165
 /note="assembly fragment:02147"

ORIGIN
 Query Match 1.1%; Score 47; DB 2: Length 110000;
 Best Local Similarity 100.0%; Pred.No. 4.1e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3829 ATGGACCTGGTACAGTCTTCGTCGACTTTCTTCAGTTGTTCTCTC 3875
 Db 101301 ATGGACCTGGTACAGTCTTCGTCGACTTTCTTCAGTTGTTCTCTC 101255
 RESULT 75
 BX004762 1/c
 WPCOMMENT
 Sequence split into 5 fragments LOCUS BX004762 Accession BX004762
 Fragment Name Begin End
 BX004762_0 1 110000
 BX004762_1 100001 210000
 BX004762_2 200001 310000
 BX004762_3 300001 410000
 BX004762_4 400001 442165
 Continuation (2 of 5) of BX004762 from base 100001 (BX004762 Mus musculus chromosome X c
 Query Match 1.1%; Score 47; DB 2: Length 110000;
 Best Local Similarity 100.0%; Pred.No. 4.1e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3829 ATGGACCTGGTACAGTCTTCGTCGACTTTCTTCAGTTGTTCTCTC 3875
 Db 1301 ATGGACCTGGTACAGTCTTCGTCGACTTTCTTCAGTTGTTCTCTC 1255
 Search completed: March 1, 2004, 18:06:08
 Job time : 15969 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2004, 12:56:07 ; Search time 9532 Seconds

(without alignments)
13439.861 Million cell updates/sec

Title: US-09-787-097-13

Perfect score: 4290

Sequence: 1 atggtggcgcagcggcggc.....agcctgggacctgcattctga 4290

Scoring table:

OLIGO NUC

GAPOP 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2368	55.2	3976	29	AY418587
2	808	18.8	1201	13	BX440935
3	681	15.9	681	12	BM783739
4	680	15.9	680	12	BM783258

5	680	15.9	680	12	BM783773
6	680	15.9	680	12	BM783773
7	677	15.8	677	12	BM783718
8	673	15.7	673	12	BM783685
9	632	14.7	632	12	BM773003
10	632	14.7	632	12	BM783264
11	612	14.3	792	14	CD644584
12	581	13.5	683	12	BM783788
13	580	13.5	633	10	BE894518
14	576	13.4	677	12	BM783821
15	556	13.0	917	10	BF672370
16	516	12.0	962	13	BX431593
17	485	11.3	486	12	BM772968
18	471	11.0	591	12	BT850289
19	435	10.1	598	10	BF855006
20	426	9.9	746	14	CB242644
21	425	9.9	595	10	BF853681
22	413	9.6	744	13	BQ183626
23	402	9.4	567	10	BF854487
24	400	9.3	544	10	BF853962
25	399	9.3	643	14	CA305924
26	390	9.1	3671	29	AY418588
27	389	9.1	739	14	CA502816
28	382	8.9	674	12	BG563729
29	369	8.6	502	9	AL048842
30	363	8.5	991	10	BE880905
31	360	8.4	484	10	BF963698
32	355	8.3	953	10	BF313142
33	353	8.2	560	10	BF853965
34	352	8.2	390	12	BG673763
35	347	8.1	681	12	BM981492
36	335	7.8	706	10	BE259981
37	333	7.8	706	12	BI089332
38	326	7.6	484	12	BT004303
39	309	7.2	399	10	BF880955
40	300	7.0	671	9	AE18764
41	299	7.0	383	9	AI372809
42	298	6.9	427	10	AM579586
43	283	6.6	320	9	AI372810
44	283	6.6	400	9	AA350293
45	281	6.6	483	10	BF853021
46	270	6.3	281	10	BG015062
47	270	6.3	396	9	AA350292
48	269	6.3	524	10	BE709733
49	253	5.9	778	13	BUS31905
50	249	5.8	589	9	AI818750
51	243	5.7	313	10	BF928756
52	241	5.6	416	14	BE7660
53	225	5.2	445	10	BE671400
54	223	5.2	464	9	AI076479
55	185	4.3	500	13	BX644466
56	176	4.1	517	14	T69367
57	171	4.0	609	10	BF206327
58	167	3.9	297	9	AA368140
59	165	3.8	265	13	BQ370569
60	165	3.8	980	13	BUI58314
61	136	3.2	355	14	T92696
62	135	3.1	413	10	BF475948
63	129	3.0	279	9	AA297214
64	126	2.9	370	9	AA771958
65	125	2.9	319	9	AA298732
66	123	2.9	230	9	AA297697
67	123	2.9	683	10	BF700035
68	119	2.8	837	28	BZ610572
69	115	2.7	134	12	BG945444
70	113	2.6	530	13	BX669805
71	110	2.6	363	29	CE218575
72	106	2.5	942	10	BF316716
73	102	2.4	496	13	BQ307981
74	97	2.3	160	10	BE171376
75	94	2.2	517	28	AQ826412
76	92	2.1	527	14	H11351
77	89	2.1	644	14	CB425680

Db	2638	TGAAATTTGTTGAGGCTACTGTGACCTTGTAGTCATTTGCTTGAGCAACAGGCTGTGGCTG	2697
Qy	3012	GTGTACTGATCCAGCAATGACTGGCAAGGAAATGCATAGAGGGTTCTTATAAAGGACC	3071
Db	2698	GTGTACTGATCCAGCAATGACTGGCAAGGAAATGCATAGAGGGTTCTTATAAAGGACC	2757
Qy	3072	AGTGAAGATGCTTTGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTC	3131
Db	2758	AGTGAAGATGCTTTGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTC	2817
Qy	3132	CAGCATGTGTAGAGACAGCAGATACACTGGTCTTTTCAATTCAGTGTCCAGTTGCCA	3191
Db	2818	CAGCATGTGTAGAGACAGCAGATACACTGGTCTTTTCAATTCAGTGTCCAGTTGCCA	2877
Qy	3192	ATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGAC	3251
Db	2878	ATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGAC	2937
Qy	3252	CACAGGCAAGCATGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGAGG	3311
Db	2938	CACAGGCAAGCATGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGAGG	2997
Qy	3312	GAATGTGAGCATGCAAGTGCATGCAATGCGGCGCTCTGTGCAACACCAACACGGGCAA	3371
Db	2998	GAATGTGAGCATGCAAGTGCATGCAATGCGGCGCTCTGTGCAACACCAACACGGGCAA	3057
Qy	3372	GTGCTTCTGCACCAACCAAGGCGCTCAAGGGGGAGCGAGTGCAGCATGTGAGGTAGAAA	3431
Db	3058	GTGCTTCTGCACCAACCAAGGCGCTCAAGGGGGAGCGAGTGCAGCATGTGAGGTAGAAA	3117
Qy	3432	TCGATACAGGAACCTCTCAGAGGAACATGTTATATACCTTCTTATTCAGTATCA	3491
Db	3118	TCGATACAGGAACCTCTCAGAGGAACATGTTATATACCTTCTTATTCAGTATCA	3177
Qy	3492	GTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGC	3551
Db	3178	GTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGC	3237
Qy	3552	TACTCTGAGCAACAAACAGGGATTTGGACATGTTTCAATGCTTCCAGGAATTTCAA	3611
Db	3238	TACTCTGAGCAACAAACAGGGATTTGGACATGTTTCAATGCTTCCAGGAATTTCAA	3297
Qy	3612	CCTCAACATACCTGGGCTGCGCTTCTCAGCTGGAACCCAGCTGGAAGAGATGCC	3671
Db	3298	CCTCAACATACCTGGGCTGCGCTTCTCAGCTGGAACCCAGCTGGAAGAGATGCC	3357
Qy	3672	TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTGA	3731
Db	3358	TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTGA	3417
Qy	3732	TTTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCAT	3791
Db	3418	TTTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCAT	3477
Qy	3792	CAAAATTCAGATTCGCTTCTCTCAGCAGCAAAATTTTATGGACCTGGTACAGTTCTCGT	3851
Db	3478	CAAAATTCAGATTCGCTTCTCTCAGCAGCAAAATTTTATGGACCTGGTACAGTTCTCGT	3537
Qy	3852	GACTTTCTCAGTGTGTTTCCCTCTCTTGTCTCTGGTGGCTGCTGTGTTTGGAAATCAA	3911
Db	3538	GACTTTCTCAGTGTGTTTCCCTCTCTTGTCTCTGGTGGCTGCTGTGTTTGGAAATCAA	3597
Qy	3912	ACAAAGTTGTTGGGCTCCAGAGCTAGAGACAACTTCTTCAGAGATGCAACAGATGGC	3971
Db	3598	ACAAAGTTGTTGGGCTCCAGAGCTAGAGACAACTTCTTCAGAGATGCAACAGATGGC	3657
Qy	3972	CAGCGCTCCCTTGTGCTCTGTAATGTCGCTTGGAAACAGATGAGAGGCTCTGATCT	4031
Db	3658	CAGCGCTCCCTTGTGCTCTGTAATGTCGCTTGGAAACAGATGAGAGGCTCTGATCT	3717
Qy	4032	TATTGGGGGAGTATAAGAGCTGTTTCCCAACCCATTCGACTGGAGCGCTGTTTGGCAA	4091

Db	3718	TATTGGGGGAGTATAAAGACTGTTCCCAAAACCCATTGCACTGGAGCGCTGTTTGGCAA	3777
Qy	4092	CAAAGCCGCTGTCTCTCTCTGTGTTTGTGAGGCTCCTCGAGGCTTGGTGGCATCCCTCC	4151
Db	3778	CAAAGCCGCTGTCTCTCTCTGTGTTTGTGAGGCTCCTCGAGGCTTGGTGGCATCCCTCC	3837
Qy	4152	TCCTGGGAGTCAGATCTTCTGTGTCGCGCAGGCGCTGTGAGCATTTCTCAGCAGATGCC	4211
Db	3838	TCCTGGGAGTCAGATCTTCTGTGTCGCGCAGGCGCTGTGAGCATTTCTCAGCAGATGCC	3897
Qy	4212	GATAGTGTACAGGAGAAAGTCAGGAGCGCTGAGAAACCGGAGCAGCAGCCCTCGCACA	4271
Db	3898	GATAGTGTACAGGAGAAAGTCAGGAGCGCTGAGAAACCGGAGCAGCAGCCCTCGCACA	3957
Qy	4272	GCCTGGGACCTGCATCTGA	4290
Db	3958	GCCTGGGACCTGCATCTGA	3976
RESULT 2			
LOCUS	EX440935	1201 bp	mRNA linear EST 15-MAY-2003
DEFINITION	EX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone		
ACCESSION	CS0DF012YL19	5-PRIME, mRNA sequence.	
VERSION	EX440935		
KEYWORDS	EX440935.1	GI:30781889	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1. (bases 1 to 1201)		
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10212.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF012CF10QPl&cluster=10212.f. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0DF012CF10QPl. Location/Qualifiers 1. 1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DF012YL19" /tissue_type="FETAL BRAIN" /dev_stage="fetal" /clone_lib="Homo sapiens FETAL BRAIN" /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
FEATURES			
source			
1. 1201			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DF012YL19"			
/tissue_type="FETAL BRAIN"			
/dev_stage="fetal"			
/clone_lib="Homo sapiens FETAL BRAIN"			
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."			
ORIGIN			
Query Match 18.8%; Score 808; DB 13; Length 1201;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 908; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	2337	GTGCATTCCTGCGCCGAAATATCTGTGGCATTTGGTGGCAATTCATG	2396
Db	57	GTGCATTCCTGCGCCGAAATATCTGTGGCATTTGGTGGCAATTCATG	116
Qy	2397	TTTGAATTTACTACTGCGCAAGAGAAATATGACATGCTAAATTTCTCTAGGAACCA	2456

Db 117 TTTGAAATTAATCTACTGCGAAGGAGAAATTATGCAATGCTAAATTTGTTCTGTAGAAACCA 176
QY 2457 CAATCCCTTTTGGCTTCTCTTACAAACCCAGAAGGTAGAAATTTGTCCTTAAGCAGCT 2516
Db 177 CAATGCCCTTTTGGCTTCTCTTACAAACCCGTGATGAAGGTAGAAATTTGTCCTTAAGCAGCT 236
QY 2517 GCGAATATGAGTCATCTCAGAGCATGTCGAAGTCACCTTAACCCATGGGTGCGCT 2576
Db 237 GCGAATATGAGTCATCTCAGAGCATGTCGAAGTCACCTTAACCCATGGGTGCGCT 296
QY 2577 TCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCCTCCCAATTTACAAATAGTTT 2636
Db 297 TCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCCTCCCAATTTACAAATAGTTT 356
QY 2637 ACTACAGTGGATGCGCTGTCGAGCCAGTCAGTGTGGAATCTGTGGAATTTTATCAGAAC 2696
Db 357 ACTACAGTGGATGCGCTGTCGAGCCAGTCAGTGTGGAATCTGTGGAATTTTATCAGAAC 416
QY 2697 CAGTACTCGGGACCTGAAGGTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2756
Db 417 CAGTACTCGGGACCTGAAGGTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 476
QY 2757 AAGGCTGCAACACACAGTGTGAAGCAGTGCAGACACCATGTGCTGTGAGCAGCATG 2816
Db 477 AAGGCTGCAACACACAGTGTGAAGCAGTGCAGACACCATGTGCTGTGAGCAGCATG 536
QY 2817 TGGAGATTGCCACGCGCAGCTCTGAGTGCATGTGTCGAGCAGCAGCATGAGCAGTGTGT 2876
Db 537 TGGAGATTGCCACGCGCAGCTCTGAGTGCATGTGTCGAGCAGCAGCATGAGCAGTGTGT 596
QY 2877 GGAATCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2936
Db 597 GGAATCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 656
QY 2937 GAGCAGCTGCCCCCTGAAATTTGTCAGCTACTGTACCTGTAGTCAATGCTTGGAGCA 2996
Db 657 GAGCAGCTGCCCCCTGAAATTTGTCAGCTACTGTACCTGTAGTCAATGCTTGGAGCA 716
QY 2997 ACCAGCTGTGCTGTCTGATCCAGCAATCTGCGCAAGGAAATGTCATGAGGG 3056
Db 717 ACCAGCTGTGCTGTCTGATCCAGCAATCTGCGCAAGGAAATGTCATGAGGG 776
QY 3057 TTCCTATAAGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGGAATTTCTATCCACA 3116
Db 777 TTCCTATAAGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGGAATTTCTATCCACA 836
QY 3117 GCCCTGTCTCAATCCAGCATGTCTAGAGCAGCAGATACCACTGGTCTTTTCAATCA 3176
Db 837 GCCCTGTCTCAATCCAGCATGTCTAGAGCAGCAGATACCACTGGTCTTTTCAATCA 896
QY 3177 CTGTCCAGCTTGCCAAATGCAACGCGCAGTAAATGTCATCAATCAGAGCATCTGTGAGAA 3236
Db 897 CTGTCCAGCTTGCCAAATGCAACGCGCAGTAAATGTCATCAATCAGAGCATCTGTGAGAA 956
QY 3237 GTGTGAGAAC 3246
Db 957 GTGTGAGAAC 966

RESULT 3
LOCUS BM783739 681 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5',
mRNA sequence.
ACCESSION BM783739
VERSION BM783739.1 GI:19131971
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: H column: 05
High quality sequence stop: 681.
FEATURES
Location/Qualifiers
1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-H05"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."

ORIGIN

Query Match 15.9%; Score 681; DB 12; Length 681;
Best/Local Similarity 100.0%; Pred. No. 0;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2512 CAGCTGCGAATTAATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGCTC 2571
Db 1 CAGCTGCGAATTAATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGCTC 60
QY 2572 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGTGGGAAGATATGTCCTCCCATTTACAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGTGGGAAGATATGTCCTCCCATTTACAAT 120
QY 2632 AGTTTACTACAGTGTGATGCCGTCTCAGCCAGTCGTGATTCGTGGATTTTATCA 2691
Db 121 AGTTTACTACAGTGTGATGCCGTCTCAGCCAGTCGTGATTCGTGGATTTTATCA 180
QY 2692 GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGT 2751
Db 181 GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGT 240
QY 2752 TGTGAAAGCCCTGCAAAACCAAGTCCTAAGCAGTCCCGGACACCATGTGCTTGAAGACA 2811
Db 241 TGTGAAAGCCCTGCAAAACCAAGTCCTAAGCAGTCCCGGACACCATGTGCTTGAAGACA 300
QY 2812 GCATGTGGAGATTGACCCAGGCGCAGCTCTGAGTCATGTGGTCAGCAACATGAAGCAG 2871
Db 301 GCATGTGGAGATTGACCCAGGCGCAGCTCTGAGTCATGTGGTCAGCAACATGAAGCAG 360
QY 2872 TGTGTGGACTCAATGCCCTATGTGCGCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGGACTCAATGCCCTATGTGCGCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
QY 2932 ACGATGAGCACCTGCTCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCATTTGCTTG 2991

```

Db 421 ACGATGACGACCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACTGTACTGTCTGTTG 480
Qy 2992 GAGCAACACGAGCTGTGGTGTGTGTACTGATCCAGCAATACTGCGCAAGGAAATGCATA 3051
Db 481 GAGCAACACGAGCTGTGGTGTGTGTACTGATCCAGCAATACTGCGCAAGGAAATGCATA 540
Qy 3052 GAGGTTTCTTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTACAGAAATTTCTAT 3111
Db 541 GAGGTTTCTTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTACAGAAATTTCTAT 600
Qy 3112 CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGCAGCAGCATACAATGCTGTTTC 3171
Db 601 CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGCAGCAGCATACAATGCTGTTTC 660
Qy 3172 ATTCACTGTCAGCTTGCCAA 3192
Db 661 ATTCACTGTCAGCTTGCCAA 681

```

RESULT 4
 BM783258
 LOCUS
 DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
 mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 680)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL
 COMMENT

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel.: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.krribb.re.kr

Plate: 31 row: H column: 08

High quality sequence stop: 680.

Location/Qualifiers

1..680

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S5SNU484-31-H08"

/sex="M"

/tissue_type="Stomach"

/cell_type="Epithelial"

/lab_host="Top10F"

/clone_lib="S5SNU484"

/note="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI;

Site: 2: NciI; The poly (A)+ RNA was decapped with tobacco

acid pyrophosphatase (TAP) and ligated with DNA-RNA linker

including EcoRI site by treatment of T4 RNA ligase. The

first strand cDNA was synthesized from oligo dt-selected

mRNA by priming with dt-tailed vector. The dt-tailed

vector was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion

of EcoRI which site is also included in vector. An RNA

strand converted to a DNA strand by Okayama-Berg method.

The obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F by electroporation

method."

ORIGIN

Query Match 15.9%; Score 680; DB 12; Length 680;
 Best Local Similarity 100.0%; Pred. No. 1.5e-313;
 Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 2512 CAGCTGCGAATATGCAATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTACAAAT 2571
Db 1 CAGCTGCGAATATGCAATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTACAAAT 60
Qy 2572 GGCCTTCGGAAGATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTACAAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTACAAAT 120
Qy 2632 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTGCTGGATTCGTGGATTTATCA 2691
Db 121 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTGCTGGATTCGTGGATTTATCA 180
Qy 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGC 2751
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGC 240
Qy 2752 TGTGAAGGCTGCGAAGATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTACAAAT 2811
Db 241 TGTGAAGGCTGCGAAGATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTACAAAT 300
Qy 2812 GCATGTGGAGATTGCAACCAAGGCTGCTGAGTGCATCTGCTGAGCAACATGAAGCAG 2871
Db 301 GCATGTGGAGATTGCAACCAAGGCTGCTGAGTGCATCTGCTGAGCAACATGAAGCAG 360
Qy 2872 TGTGTCGATCCCAATGCCTATGTGGCTCTCTCCCTTTTGGCAGTGTATGGAATGCTAT 2931
Db 361 TGTGTCGATCCCAATGCCTATGTGGCTCTCTCCCTTTTGGCAGTGTATGGAATGCTAT 420
Qy 2932 ACATGAGCAGCTGCCCCCTGAAATTTCTGAGGCTACTGTACCTGTAGTCAATGCTTGG 2991
Db 421 ACATGAGCAGCTGCCCCCTGAAATTTCTGAGGCTACTGTACCTGTAGTCAATGCTTGG 480
Qy 2992 GAGCAACCAAGGCTGTGGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA 3051
Db 481 GAGCAACCAAGGCTGTGGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA 540
Qy 3052 GAGGTTTCTTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTACAGAAATTTCTAT 3111
Db 541 GAGGTTTCTTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTACAGAAATTTCTAT 600
Qy 3112 CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGCAGCAGCATACAATGCTGTTTC 3171
Db 601 CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGCAGCAGCATACAATGCTGTTTC 660
Qy 3172 ATTCACTGTCAGCTTGCCAA 3191
Db 661 ATTCACTGTCAGCTTGCCAA 680

```

RESULT 5

BM783773

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

BM783773

K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',

mRNA sequence.

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

BM783773 680 bp mRNA linear EST 05-MAR-2002
 K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
 mRNA sequence.

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

BM783773 1 GI:19132005

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 680)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krribb.re.kr
Plate: 36 row: D column: 02
High quality sequence scop: 680.
Location/Qualifiers

FEATURES

source

1..680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-D02"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation
method."

ORIGIN

Query Match 15.9%; Score 680; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 1.5e-313;
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATTAATGCACTGATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGGTC 2571
DB 1 CAGCTGCGAATTAATGCACTGATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGCTGCTGGGAAGATATGTCCCATTTACAAAT 2631
DB 61 GGCCTTCGGAAGATCAATGTGCTCTACTGCTGCTGGGAAGATATGTCCCATTTACAAAT 120

QY 2632 AGTTTACTACAGTGGATCGCTCTGAGCCAGTGTGATCTGGAATTTTATCA 2691
DB 121 AGTTTACTACAGTGGATCGCTCTGAGCCAGTGTGATCTGGAATTTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTGC 2751
DB 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTGC 240

QY 2752 TGTGAAGGCTGCAACCCAGTGTGAGCAGTGGCGGACACCATGTGCTTGAAGACA 2811
DB 241 TGTGAAGGCTGCAACCCAGTGTGAGCAGTGGCGGACACCATGTGCTTGAAGACA 300

QY 2812 GCATGTGAGATTTGACACAGGGGAGCTCTGAGTGCATGTGCTGATGCTGCAACATGAAGCAG 2871
DB 301 GCATGTGAGATTTGACACAGGGGAGCTCTGAGTGCATGTGCTGATGCTGCAACATGAAGCAG 360

QY 2872 TGTGTGAGTCAATGCTATGTGGCTCTCTTCCTTTTGGCCAGTGTATGGAATGGTAT 2931
DB 361 TGTGTGAGTCAATGCTATGTGGCTCTCTTCCTTTTGGCCAGTGTATGGAATGGTAT 420

QY 2932 ACGATGAGCAGTCCGCCCTGAAATTTGTCAGGCTACTGTACTCTAGTCTATTGCTTG 2991
DB 421 ACGATGAGCAGTCCGCCCTGAAATTTGTCAGGCTACTGTACTCTAGTCTATTGCTTG 480

QY 2992 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGCCAAGGGAATGCATA 3051

DB 481 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATATCCAGGAATATGCATA 540

QY 3052 GAGGTTTCTATAAAGGACCACTGAGATGCTTCCGAAGCCCTTACAGGAATTTCTAT 3111
DB 541 GAGGTTTCTATAAAGGACCACTGAGATGCTTCCGAAGCCCTTACAGGAATTTCTAT 600

QY 3112 CCACAGCCCTGCTCAATTCACAGCATGCTCTAGAGGACAGATACACTGCTTCTTC 3171
DB 601 CCACAGCCCTGCTCAATTCACAGCATGCTCTAGAGGACAGATACACTGCTTCTTC 660

QY 3172 ATTCAGTCTCCAGCTTGCCA 3191
DB 661 ATTCAGTCTCCAGCTTGCCA 680

RESULT 6
BG678679 933 bp mRNA linear EST 01-MAY-2001
LOCUS 60262447F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
DEFINITION mRNA sequence.
ACCESSION BG678679
VERSION BG678679.1 GI:13910076
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10602 row: j column: 20
High quality sequence scop: 782.
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 15.9%; Score 680; DB 12; Length 933;
Best Local Similarity 99.9%; Pred. No. 1.5e-313;
Matches 730; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2997 ACCAGGCTGTGGCTGTGCTACTGATCCAGCAATATCCAGGAATATGCATAGAGG 3056
DB 2 ACCAGGCTGTGGCTGTGCTACTGATCCAGCAATATCCAGGAATATGCATAGAGG 61

QY 3057 TTCCTATAAAGGACCACTGAGATGCTTCCGAAGCCCTTACAGGAATTTCTATCCACA 3116
DB 62 TTCCTATAAAGGACCACTGAGATGCTTCCGAAGCCCTTACAGGAATTTCTATCCACA 121

QY 3117 GCCCTGTCTCAATTCAGCATGCTCTAGAGGACAGATACACTGCTTCTTCATTC 3176
DB 122 GCCCTGTCTCAATTCAGCATGCTCTAGAGGACAGATACACTGCTTCTTCATTC 181

```
QY 3177 CTGTCCAGCTTGCATATGCAACGGCCACAGTAAATGATGCAATCAATCAGAGCATCTGTGAGAA 3236
Db 182 CTGTCCAGCTTGCATATGCAACGGCCACAGTAAATGATGCAATCAATCAGAGCATCTGTGAGAA 241
QY 3237 GTGTGAGAACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3296
Db 242 GTGTGAGAACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 301
QY 3297 TCCACCAATGTGAGGGAAATGTGAGCCATGCAAGTGAATGGGCAAGCGTCTCTGTGCAAA 3356
Db 302 TCCACCAATGTGAGGGAAATGTGAGCCATGCAAGTGAATGGGCAAGCGTCTCTGTGCAAA 361
QY 3357 CACCAACACAGGGCAAGTGTCTTCTGACACCAAGGGGGTCAAGGGGGGACAGTCCAGCT 3416
Db 362 CACCAACACAGGGCAAGTGTCTTCTGACACCAAGGGGGTCAAGGGGGGACAGTCCAGCT 421
QY 3417 ATGTGAGGTGAAATCGATACCAAGGAAACCCCTCTCAGAGGACATGTTATTATCTCT 3476
Db 422 ATGTGAGGTGAAATCGATACCAAGGAAACCCCTCTCAGAGGACATGTTATTATCTCT 481
QY 3477 TCTTATTGACTATCAGTTCACCTTTACTCTATCCAGAGATGATCGCTATTACACAGC 3536
Db 482 TCTTATTGACTATCAGTTCACCTTTACTCTATCCAGAGATGATCGCTATTACACAGC 541
QY 3537 TATCAATTTTGTGCTACTCTGACGAAACAAACAGGGATTTGGACATGTTCTATCATGC 3596
Db 542 TATCAATTTTGTGCTACTCTGACGAAACAAACAGGGATTTGGACATGTTCTATCATGC 601
QY 3597 CTCCAAGAAATTAACCTCAACATCAGCTGGCTGGCCAGTTTCTCAGCTGGAACCCAGGC 3656
Db 602 CTCCAAGAAATTAACCTCAACATCAGCTGGCTGGCCAGTTTCTCAGCTGGAACCCAGGC 661
QY 3657 TGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGAGATGACAAAGATAGTTTCTC 3716
Db 562 TGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGAGATGACAAAGATAGTTTCTC 721
QY 3717 TAATGAGAAGT 3727
Db 722 TAATGAGAAGT 732

RESULT 7
BM783718 677 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',
DEFINITION mRNA sequence.
ACCESSION BM783718
VERSION BM783718.1 GI:19131950
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.O., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoseun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: F column: 01
High quality sequence stop: 677.
Location/Qualifiers
1. .677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
```

```
/clone="S5SNU484-34-F01"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_lines="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ189Pl; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."
```

ORIGIN

```
Query Match 15.8%; Score 677; DB 12; Length 677;
Best Local Similarity 100.0%; Pred.No. 3.1e-312;
Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2512 CAGCTGCGAATTAATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC 2571
Db 1 CAGCTGCGAATTAATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC 60
QY 2572 GGCCTTCGGAAGATCAATGTGTCTTACTGTGGTCTGGGAAGATATGTCCCATTTTCAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGTCTTACTGTGGTCTGGGAAGATATGTCCCATTTTCAAT 120
QY 2632 AGTTTACTACGTGATGCGCTGAGCCAGTGTGATGCTGTGGATTTCTGTGGATTTTATCA 2691
Db 121 AGTTTACTACGTGATGCGCTGAGCCAGTGTGATGCTGTGGATTTCTGTGGATTTTATCA 180
QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCT 2751
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCT 240
QY 2752 TGTGAAGGCTTCGAAACCAACAGTGTGTAAGCAGTCCGGACACCATGTGCCTTGAGGACA 2811
Db 241 TGTGAAGGCTTCGAAACCAACAGTGTGTAAGCAGTCCGGACACCATGTGCCTTGAGGACA 300
QY 2812 GCATGTGGAGATGTGACACGCGGAGCTGTGATGTCATGTGGTGACGAAACATGAAGCAG 2871
Db 301 GCATGTGGAGATGTGACACGCGGAGCTGTGATGTCATGTGGTGACGAAACATGAAGCAG 360
QY 2872 TGTGTGGACTCCAATGCTTATGTGCGCTCTTCCCTTTTGGCCAGTGTGTAATGGTAT 2931
Db 361 TGTGTGGACTCCAATGCTTATGTGCGCTCTTCCCTTTTGGCCAGTGTGTAATGGTAT 420
QY 2932 AGCATGAGCACCTGCCCCCTGAAAATTTCTCAGGCTACTGTACCTGTAGTCTGATGCTTG 2991
Db 421 AGCATGAGCACCTGCCCCCTGAAAATTTCTCAGGCTACTGTACCTGTAGTCTGATGCTTG 480
QY 2992 GAGCAACCAAGGCTGTGGTGTGTACTGATCCAGCAATATCTGGCAAGGGGAAATGCATA 3051
Db 481 GAGCAACCAAGGCTGTGGTGTGTACTGATCCAGCAATATCTGGCAAGGGGAAATGCATA 540
QY 3052 GAGGTTTCTTATAAGGACCAAGTGAAGTGCCTTCGACGCCCCCTACAGGAATTTCTAT 3111
Db 541 GAGGTTTCTTATAAGGACCAAGTGAAGTGCCTTCGACGCCCCCTACAGGAATTTCTAT 600
QY 3112 CCACAGCCCTCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACCTGGTCTTTC 3171
Db 601 CCACAGCCCTCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACCTGGTCTTTC 660
QY 3172 ATTCACTGTCCAGCTTG 3188
Db 661 ATTCACTGTCCAGCTTG 677
```

RESULT 8	BM783685	673 bp	mRNA	linear	EST 05-MAR-2002									
LOCUS	K-EST0061659	S5SNU484	Homo sapiens	CDNA clone	S5SNU484-34-B06 5',									
DEFINITION	mRNA sequence.													
ACCESSION	BM783685	1	GI:19131917											
KEYWORDS	EST													
SOURCE	Homo sapiens	(human)												
ORGANISM	Homo sapiens													
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.													
AUTHORS	1 (bases 1 to 673) Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.													
TITLE	21C Frontier Korean EST Project 2001													
JOURNAL	Unpublished (2002)													
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 34 row: B column: 06 High quality sequence stop: 673.													
FEATURES	Location/Qualifiers													
source	1..673													
	/organism="Homo sapiens"													
	/mol_type="mRNA"													
	/db_xref="taxon:9606"													
	/clone="S5SNU484-34-B06"													
	/sex="M"													
	/tissue_type="Stomach"													
	/cell_type="Epithelial"													
	/lab_host="Top10F"													
	/clone_lib="S5SNU484"													
	/note="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI; Site 2: NotI; The poly (A) + RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."													
ORIGIN														
Query Match	15.7%; Score 673; DB 12; Length 673;													
Best Local Similarity	100.0%; Pred. No. 2.5e-310;													
Matches 673;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;													
QY	2512	CAGCTGCGAATATGCACTATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC	2571											
Db	1	CAGCTGCGAATATGCACTATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC	60											
QY	2572	GGCCTTCGGAAGATCAATGTCTCTACTGGTGGGAAGATGTCCCATTTACAAAT	2631											
Db	61	GGCCTTCGGAAGATCAATGTCTCTACTGGTGGGAAGATGTCCCATTTACAAAT	120											
QY	2632	AGTTTACTACAGTGGATGCCCTCTGAGCCAGTCAGTGTCTGTGAATTTATCA	2691											
Db	121	AGTTTACTACAGTGGATGCCCTCTGAGCCAGTCAGTGTCTGTGAATTTATCA	180											
QY	2692	GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC	2751											
Db	181	GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC	240											
QY	2752	TGTGAAGGCTGCAAAACACACAGTCTTAAGCAGTCCCGACACCATGTGCTTGAAGACA	2811											
Db	241	TGTGAAGGCTGCAAAACACACAGTCTTAAGCAGTCCCGACACCATGTGCTTGAAGACA	300											
QY	2812	GCATGTGGAGATTGCACACAGCGGACGCTCTGAGTCATGTGGTCAGCAACATGAAGCAG	2871											
Db	301	GCATGTGGAGATTGCACACAGCGGACGCTCTGAGTCATGTGGTCAGCAACATGAAGCAG	360											
QY	2872	TGTGTGGACTCAATGCTTATGTGGCTCCCTTCCCTTTGGCCAGTGTATGATGAT	2931											
Db	361	TGTGTGGACTCAATGCTTATGTGGCTCCCTTCCCTTTGGCCAGTGTATGATGAT	420											
QY	2932	ACGATGAGCAGCTGCCCCCTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTTG	2991											
Db	421	ACGATGAGCAGCTGCCCCCTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTTG	480											
QY	2992	GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATATCGCAAGGGAATGCATA	3051											
Db	481	GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATATCGCAAGGGAATGCATA	540											
QY	3052	GAGGTTTCTATAAAGCACCAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTAT	3111											
Db	541	GAGGTTTCTATAAAGCACCAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTAT	600											
QY	3112	CCACAGCCCTGCTCAATCCAGCATGTGTCTAGAGGACAGCATACACTGTCTTTTC	3171											
Db	601	CCACAGCCCTGCTCAATCCAGCATGTGTCTAGAGGACAGCATACACTGTCTTTTC	660											
QY	3172	ATTCACTGTCCAG	3184											
Db	661	ATTCACTGTCCAG	673											
RESULT 9	BM773003	632 bp	mRNA	linear	EST 04-MAR-2002									
LOCUS	K-EST0057379	S5SNU484s1	Homo sapiens	CDNA clone	S5SNU484s1-3-E09									
DEFINITION	5', mRNA sequence.													
ACCESSION	BM773003	1	GI:19102618											
KEYWORDS	EST													
SOURCE	Homo sapiens	(human)												
ORGANISM	Homo sapiens													
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.													
AUTHORS	1 (bases 1 to 632) Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.													
TITLE	21C Frontier Korean EST Project 2001													
JOURNAL	Unpublished (2002)													
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 3 row: E column: 09 High quality sequence stop: 632.													
FEATURES	Location/Qualifiers													
source	1..632													
	/organism="Homo sapiens"													
	/mol_type="mRNA"													
	/db_xref="taxon:9606"													
	/clone="S5SNU484s1-3-E09"													
	/sex="M"													
	/tissue_type="Stomach"													
	/cell_type="Epithelial"													
	/lab_host="SNU-484"													
	/clone_lib="S5SNU484"													
	/note="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI; Site 2: NotI; The poly (A) + RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."													

```
/lab_host="Top10F"
/clone_lib="S5SNU484s1"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
After analyzing and sequencing about 2,000 - 3,000
colonies in original cDNA library, the abundant cDNAs were
selected and amplified by PCR reaction using vector region
primer including T7 promoter as 5' primer and N(dn)14 as
3' primer. The PCR products were used as template for
synthesis of biotinylated single stranded RNA by in vitro
transcription reaction. The synthesized RNA probes were
hybridized with antisense single stranded cDNAs prepared
from original library and incubated with avidin-gel.
After removing DNA-RNA hybrids by centrifuge, the
subtracted cDNA libraries were constructed by
transformation of the remaining DNA into competent cells E.
coli Top10F' with electroporation method."
```

ORIGIN

Query Match 14.7%; Score 632; DB 12; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.1e-290;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2512 CAGCTCGGAATATGCAATGCTCATCTCAGAGCATGTCAGAGCTCACCTTAACCCCATGGTCT 2571
Db 1 CAGCTCGGAATATGCAATGCTCATCTCAGAGCATGTCAGAGCTCACCTTAACCCCATGGTCT 60

Qy 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGCGGAAGATATGCCCCATTTACAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGCGGAAGATATGCCCCATTTACAAT 120

Qy 2632 AGTTTACTACAGTGATGCGCTGTGAGCCAGTGATGCTGGATTCTGTGGAAATTTATCA 2691
Db 121 AGTTTACTACAGTGATGCGCTGTGAGCCAGTGATGCTGGATTCTGTGGAAATTTATCA 180

Qy 2692 GAACCCAGTACTCGGAGCTGAAGCTGCAACTGTCATCAACCCATCAATGTAGTGTC 2751
Db 181 GAACCCAGTACTCGGAGCTGAAGCTGCAACTGTCATCAACCCATCAATGTAGTGTC 240

Qy 2752 TGTGAAAGGCTGCAAAACCAAGTCTAAGCAGTGCAGACACCATGTGCTTGGAGCA 2811
Db 241 TGTGAAAGGCTGCAAAACCAAGTCTAAGCAGTGCAGACACCATGTGCTTGGAGCA 300

Qy 2812 GCATGTGGAGATTGCACAGCGGAGCTCTGAGTGCAATGCTGTCAGACATGAGCAG 2871
Db 301 GCATGTGGAGATTGCACAGCGGAGCTCTGAGTGCAATGCTGTCAGACATGAGCAG 360

Qy 2872 TGTGTGGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy 2932 ACGATGAGCACTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACGATGAGCACTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480

Qy 2992 GAGCAACAGGCTGTGCTGTGTACTGATCCAGCAATACTGCGAAAGGGAATGCATA 3051
Db 481 GAGCAACAGGCTGTGCTGTGTACTGATCCAGCAATACTGCGAAAGGGAATGCATA 540

Qy 3052 GAGGGTTTCTATAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 3111
Db 541 GAGGGTTTCTATAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 600

Qy 3112 CCACAGCCCTCTCTCAATTCAGATTCAGCATGTGCT 3143
Db 601 CCACAGCCCTCTCTCAATTCAGATTCAGCATGTGCT 632

RESULT 10
LOCUS BM783264
DEFINITION K-EST0061166 S5SNU484 Homo sapiens cDNA clone S5SNU484-32-A05 5',
linear mRNA 632 bp
BM783264
ACCESSION BM783264
VERSION 1
KEYWORDS EST, GI:19131496
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 32 row: A column: 05
High quality sequence stop: 632.
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-32-A05"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation
method."

ORIGIN

Query Match 14.7%; Score 632; DB 12; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.1e-290;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2512 CAGCTCGGAATATGCAATGCTCATCTCAGAGCATGTCAGAGCTCACCTTAACCCCATGGTCT 2571
Db 1 CAGCTCGGAATATGCAATGCTCATCTCAGAGCATGTCAGAGCTCACCTTAACCCCATGGTCT 60

Qy 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGCGGAAGATATGCCCCATTTACAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGCGGAAGATATGCCCCATTTACAAT 120

Qy 2632 AGTTTACTACAGTGATGCGCTGTGAGCCAGTGATGCTGGATTCTGTGGAAATTTATCA 2691
Db 121 AGTTTACTACAGTGATGCGCTGTGAGCCAGTGATGCTGGATTCTGTGGAAATTTATCA 180

Qy 2692 GAACCCAGTACTCGGAGCTGAAGCTGCAACTGTCATCAACCCATCAATGTAGTGTC 2751
Db 181 GAACCCAGTACTCGGAGCTGAAGCTGCAACTGTCATCAACCCATCAATGTAGTGTC 240

Qy 2752 TGTGAAAGGCTGCAAAACCAAGTCTAAGCAGTGCAGACACCATGTGCTTGGAGCA 2811
Db 241 TGTGAAAGGCTGCAAAACCAAGTCTAAGCAGTGCAGACACCATGTGCTTGGAGCA 300

Qy 2812 GCATGTGGAGATTGCACAGCGGAGCTCTGAGTGCAATGCTGTCAGACATGAGCAG 2871
Db 301 GCATGTGGAGATTGCACAGCGGAGCTCTGAGTGCAATGCTGTCAGACATGAGCAG 360

Qy 2872 TGTGTGGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy 2932 ACGATGAGCACTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACGATGAGCACTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480

Qy 2992 GAGCAACAGGCTGTGCTGTGTACTGATCCAGCAATACTGCGAAAGGGAATGCATA 3051
Db 481 GAGCAACAGGCTGTGCTGTGTACTGATCCAGCAATACTGCGAAAGGGAATGCATA 540

Qy 3052 GAGGGTTTCTATAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 3111
Db 541 GAGGGTTTCTATAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 600


```

Db      492  TTCTGTACCGCTAACACAGTCATTTGGACTCGAGAGGATATTTCACTTAAAGTCC 551
Qy      1019 CCAGAGCATCTCATTAAGCTGTGTGTCATGGAACATTTATGTGGGTTTGGAGGATATA 1078
Db      552  CCAGAGCATCTCATTAAGCTGTGTGTCATGGAACATTTATGTGGGTTTGGAGGATATA 611
Qy      1079 TGTTCAACCATCAGATTAACATGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGC 1138
Db      612  TGTTCAACCATCAGATTAACATGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGC 671
Qy      1139 TTC 1141
Db      672  TTC 674

RESULT 12
BM783788
LOCUS   K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
DEFINITION mRNA sequence.
ACCESSION BM783788
VERSION   1
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 683)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 36 row: E column: 09
           High quality sequence stop: 683.
           Location/Qualifiers
             1..683
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="S5SNU484-36-E09"
               /sex="M"
               /tissue_type="Stomach"
               /cell_type="Epithelial"
               /cell_line="SNU-484"
               /lab_host="Top10P"
               /clone_lib="S5SNU484"
               /note="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI;
               Site 2: NotI; The poly (A)+ RNA was decapped with tabacco
               acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
               including EcoRI site by treatment of T4 RNA ligase. The
               first strand cDNA was synthesized from oligo dt-selected
               mRNA by priming with dt-tailed vector. The dt-tailed
               vector was adjusted to have about 60nt. The cDNA vector
               was circularized with E. coli DNA ligase after digestion
               of EcoRI which site is also included in vector. An RNA
               strand converted to a DNA strand by Okayama-Berg method.
               The obtained cDNA vectors were used for transformation of
               competent cells E. coli Top10P by electroporation
               method."
           13.5%; Score 581; DB 12; Length 683;
           Best Local Similarity 99.7%; Pred. No. 2.9e-266;
           Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN
Query Match
Best Local Similarity
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      2512 CAGCTCGGAATATATGAGTCATCTCAGAGCATGTCTCAAGCTCACCTTAACCCCATGGGTC 2571
Db      1    CAGCTCGGAATATATGAGTCATCTCAGAGCATGTCTCAAGCTCACCTTAACCCCATGGGTC 60
Qy      2572 GGCCTTCGGAAGATCAATGTGTCTTACTGTCTCTGGAGAGATATGTCCCATTTTCAAAAT 2631
Db      61  GGCCTTCGGAAGATCAATGTGTCTTACTGTCTCTGGAGAGATATGTCCCATTTTCAAAAT 120
Qy      2632 AGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGTGGATTCTGTGGAATTTATCA 2691
Db      121  AGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGTGGATTCTGTGGAATTTATCA 180
Qy      2692 GAACCCAGTACTCGGGGACTGAAGCTGCACCTGCATCAACCCACTCAATGGTAGTGTTC 2751
Db      181  GAACCCAGTACTCGGGGACTGAAGCTGCACCTGCATCAACCCACTCAATGGTAGTGTTC 240
Qy      2752 TGTGAAGGCGCTTCCAAACACACAGTGTCTAAGCAGTGTCCGGACACCATGTGCTTTGAGGACA 2811
Db      241  TGTGAAGGCGCTTCCAAACACACAGTGTCTAAGCAGTGTCCGGACACCATGTGCTTTGAGGACA 300
Qy      2812 GCATGTGGAGATTCACACAGCGGCGAGCTCTGAGTGCATGTGTGAGCAGACATGAAGCAG 2871
Db      301  GCATGTGGAGATTCACACAGCGGCGAGCTCTGAGTGCATGTGTGAGCAGACATGAAGCAG 360
Qy      2872 TGTGTGCACTCCAAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGAAATGGTAT 2931
Db      361  TGTGTGCACTCCAAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGAAATGGTAT 420
Qy      2932 AGCATGAGCACCTGCGCCCTGAAATTTGTCAGGCTACTGTACTGTAGTGTGCTTGTG 2991
Db      421  AGCATGAGCACCTGCGCCCTGAAATTTGTCAGGCTACTGTACTGTAGTGTGCTTGTG 480
Qy      2992 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATACCTGCGCAAGGGAATGTCATA 3051
Db      481  GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATACCTGCGCAAGGGAATGTCATA 540
Qy      3052 GAGGGTTTCTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTAT 3111
Db      541  GAGGGTTTCTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTAT 600
Qy      3112 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGGAGACAGCAGATACACCTGGCTTTTC 3171
Db      601  CCACAGCCCTGCTCAATTCAGCATGTGTCTAGGAGACAGCAGATACACCTGGCTTTTC 660
Qy      3172 ATTCACTGTCCAGCTTGCCAATG 3194
Db      661  ATTCACTGTCCAGCTTGCCAATG 683

```

```

RESULT 13
BM783788
LOCUS   601433126F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3918271 5',
DEFINITION mRNA sequence.
ACCESSION BE894518
VERSION   BE894518.1 GI:10356969
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 633)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: ATCC/DCTD/DTF
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: U94M9746 row: c column: 08

High quality sequence stop: 632.

Location/Qualifiers

1. .633

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3918271"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 2 kb. Library constructed by Life

Technologies."

FEATURES

Query Match 13.5%; Score 580; DB 10; Length 633;
Best Local Similarity 100.0%; Pred. No. 8.5e-266;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3162	CTGGTCTTTTCATCTCACTGTCAGCTTGCCAAATGCAAGCGGCACAGTAAATGCAATCA	3221
Db	5	CTGGTCTTTTCATCTCACTGTCAGCTTGCCAAATGCAAGCGGCACAGTAAATGCAATCA	64
QY	3222	GAGCATCTGTGAGAAGTGTGAGAACCTGTGACACAGGCAAGCACTGCAGACCTGCATATC	3281
Db	65	GAGCATCTGTGAGAAGTGTGAGAACCTGTGACACAGGCAAGCACTGCAGACCTGCATATC	124
QY	3282	TGGCTTCTACGGTATCCCAATGAGGAAATGTCAGCCATGCAAGTGCATGGGCA	3341
Db	125	TGGCTTCTACGGTATCCCAATGAGGAAATGTCAGCCATGCAAGTGCATGGGCA	184
QY	3342	CGCGTCTCTGTGCAACCAACAGGCAAGTGTCTTGACACCAACCAAGGCGCTCAAGGG	3401
Db	195	CGCGTCTCTGTGCAACCAACAGGCAAGTGTCTTGACACCAACCAAGGCGCTCAAGGG	244
QY	3402	GGACGAGTGCACGCTATGTAGGTAGGTAAGTAATCGATACAGGAAACCTCTCAGAGGAAC	3461
Db	245	GGACGAGTGCACGCTATGTAGGTAGGTAAGTAATCGATACAGGAAACCTCTCAGAGGAAC	304
QY	3462	ATGTTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCGAAGATGA	3521
Db	305	ATGTTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCGAAGATGA	364
QY	3522	TCGCTATTACACAGCTATCAATTTTGGCTACTCCTGACGAAACAAACAGGGATTGGA	3581
Db	365	TCGCTATTACACAGCTATCAATTTTGGCTACTCCTGACGAAACAAACAGGGATTGGA	424
QY	3582	CATGTTTCATCAATGCTCCAGAAATTTCAACCTCAAGATCACCTGGGCTGCCAGTTTCTC	3641
Db	425	CATGTTTCATCAATGCTCCAGAAATTTCAACCTCAAGATCACCTGGGCTGCCAGTTTCTC	484
QY	3642	AGCTGGAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTA	3701
Db	485	AGCTGGAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTA	544
QY	3702	CAAGATAGTTTCTTAATGAGAGTTTGATTTTCGCAAC	3741
Db	545	CAAGATAGTTTCTTAATGAGAGTTTGATTTTCGCAAC	584

RESULT 14
LOCUS BM783821
DEFINITION K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5',
mRNA sequence.
ACCESSION BM783821
VERSION BM783821.1 GI:19132053
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 677)

AUTHORS

Kim N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.

TITLE

21C Frontier Korean EST Project 2001

JOURNAL

Unpublished (2002)

COMMENT

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Soeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 36 row: H column: 09

High quality sequence stop: 677.

FEATURES

source

1. .677

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S5SNU484-36-H09"

/sex="M"

/tissue_type="Stomach"

/cell_type="Epithelial"

/cell_line="SNU-484"

/lab_host="Top10F"

/clone_lib="S5SNU484"

/note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI;

Site 2: NotI; The poly (A) + RNA was decapped with tobacco

acid pyrophosphatase (TAP) and ligated with DNA-RNA linker

including EcoRI site by treatment of T4 RNA ligase. The

first strand cDNA was synthesized from oligo dr-selected

mRNA by priming with dr-tailed vector. The dr-tailed

vector was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion

of EcoRI which site is also included in vector. An RNA

strand converted to a DNA strand by Okayama-Berg method.

The obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F by electroporation

method."

ORIGIN

Query Match	13.4%;	Score 576;	DB 12;	Length 677;
Best Local Similarity	100.0%;	Pred. No. 7.1e-264;	Mismatches 0;	Indels 0; Gaps 0;
Matches 576;	Conservative 0;			
QY	2512	CAGCTGCGAATAATCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC	2571	
Db	1	CAGCTGCGAATAATCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC	60	
QY	2572	GGCCTTCGGAAGATCAATGTCTCTACTGGTCTGGGAGATATGTCCTCCCAATTTACAAT	2631	
Db	61	GGCCTTCGGAAGATCAATGTCTCTACTGGTCTGGGAGATATGTCCTCCCAATTTACAAT	120	
QY	2632	AGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGTGCTGGATTCTGGAATTTATCA	2691	
Db	121	AGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGTGCTGGATTCTGGAATTTATCA	180	
QY	2692	GAAACCCAGTACTCGGGGACTGAAGCTGCACCTGCATCAACCCCACTCAATGGTAGTGC	2751	
Db	181	GAAACCCAGTACTCGGGGACTGAAGCTGCACCTGCATCAACCCCACTCAATGGTAGTGC	240	
QY	2752	TGTGAAAGCCCTGCAAAACCAACAGTCTAAGCACTGCCGACACCACTGTCCTTGAAGGACA	2811	
Db	241	TGTGAAAGCCCTGCAAAACCAACAGTCTAAGCACTGCCGACACCACTGTCCTTGAAGGACA	300	
QY	2812	GCATGTGGAGATTGCAACGAGCGGAGCTCTGAGTGCAATGCGTGCAGCAACATGAGGAG	2871	
Db	301	GCATGTGGAGATTGCAACGAGCGGAGCTCTGAGTGCAATGCGTGCAGCAACATGAGGAG	360	

```

QY 2872 TGTTGGACTCCAATGCTATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTTGGACTCCAATGCTATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

QY 2932 ACGATGAGCACTGCCCTCCCTGAAATTTGTTCAAGGTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACGATGAGCACTGCCCTCCCTGAAATTTGTTCAAGGTACTGTACCTGTAGTCAATGCTTG 480

QY 2992 GAGCAACACAGCTGTGCTCGTGTACTGATCCAGCAATACTGCAAGGGAATGCATA 3051
Db 481 GAGCAACACAGCTGTGCTCGTGTACTGATCCAGCAATACTGCAAGGGAATGCATA 540

QY 3052 GAGGTTTCTTATAAGGACAGTGAAGATCCTTGG 3087
Db 541 GAGGTTTCTTATAAGGACAGTGAAGATCCTTGG 576

RESULT 15
BF672370 917 bp mRNA linear EST 21-DEC-2000
LOCUS 602150605F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4291680 5',
DEFINITION mRNA sequence.
ACCESSION BF672370
VERSION BF672370.1 GI:11946265
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1136 row: j column: 01
High quality sequence stop: 563.
FEATURES
    source
        1. 917
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4291680"
                /lab_hosts="DH10B (TI phage-resistant)"
                /clone_lib="NIH_MGC_81"
                /notes="Organ: muscle (skeletal); Vector: pDNR-LIB
                (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
                (ggccattatggc); 5' and 3' adaptors were used in cloning
                as follows: 5' adaptor sequence: 5'-CAGGCGCAATTATGGCC-3'
                and 3' adaptor sequence:
                5'-ATTCTAGAGCCGCGGCGGCACATG-dt(30)BN-3' (where B = A,
                C, or G and N = A, C, G, or T). Average insert size
                1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
                inserts by PCR. This library was enriched for full-length
                clones and was constructed by Clontech Laboratories (Palo
                Alto, CA)."

Query Match 13.0%; Score 556; DB 10; Length 917;
Best Local Similarity 100.0%; Pred. NO. 2.9e-254;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2833 GGCAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGTGCACTCCAATGCTAT 2892
Db 1 GGCAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGTGCACTCCAATGCTAT 60

```

```

QY 2893 GTGGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCCT 2952
Db 61 GTGGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCCT 120

QY 2953 GAAAATTTGTTCAAGGTACTGTACCTGTAGTCAATTCCTTTGAGAGCAACCAAGGTGTGGCTGG 3012
Db 121 GAAAATTTGTTCAAGGTACTGTACCTGTAGTCAATTCCTTTGAGAGCAACCAAGGTGTGGCTGG 180

QY 3013 TGTACTGATCCAGCAATACTGCAAGGGAATGCATAGAGGGTTCCTATAAAGGACCA 3072
Db 181 TGTACTGATCCAGCAATACTGCAAGGGAATGCATAGAGGGTTCCTATAAAGGACCA 240

QY 3073 GTGAAGATGCTTTCGCAAGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCC 3132
Db 241 GTGAAGATGCTTTCGCAAGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCC 300

QY 3133 AGCATGTGCTTAGAGGACAGCAGATACAACCTGGTCTTTCTACTGTCAGCTTGCCAA 3192
Db 301 AGCATGTGCTTAGAGGACAGCAGATACAACCTGGTCTTTCTACTGTCAGCTTGCCAA 360

QY 3193 TGCAACGGCCACAGTAAATGATCAATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 3252
Db 361 TGCAACGGCCACAGTAAATGATCAATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 420

QY 3253 ACAGCAAGCACTGCAGACCTGCATATCTGCTTCTAGGTGATCCACCACCAATGAGGG 3312
Db 421 ACAGCAAGCACTGCAGACCTGCATATCTGCTTCTAGGTGATCCACCACCAATGAGGG 480

QY 3313 AAATGTGAGCCATGCAAGTGAATGGGACCGCTCTCTGTGCAACACCAACACGGGCAAG 3372
Db 481 AAATGTGAGCCATGCAAGTGAATGGGACCGCTCTCTGTGCAACACCAACACGGGCAAG 540

QY 3373 TGCCTTCGACCAACCA 3388
Db 541 TGCCTTCGACCAACCA 556

RESULT 16
BX431593 962 bp mRNA linear EST 15-MAY-2003
LOCUS CS0DF012YL19 5-PRIME, mRNA sequence.
DEFINITION BX431593
ACCESSION BX431593
VERSION BX431593.1 GI:30781048
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secret@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10212.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG017B02_CS01546_1&cluster=10212.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG017B02_CS01546_1.
FEATURES
    source
        1. 962
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DF012YL19"
                /tissue_type="FETAL BRAIN"
                /dev_stage="fetal"

```

ORIGIN

Query Match 12.0%; Score 516; DB 13; Length 962;
Best Local Similarity 99.7%; Pred. No. 4.1e-235;
Matches 616; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3159 CAACCTGGCTTTCATTCAGTCTGCAGCTTGCATGCAAGGCGGACACAGTAAATGCATCAA 3218
Db CAACCTGGCTTTCATTCAGTCTGCAGCTTGCATGCAAGGCGGACACAGTAAATGCATCAA 111

3219 TCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACCTGCAT 3278
Db TCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACCTGCAT 171

3279 ATCTGGCTTCTACGCTGATCCCAATGAGGGAATGTGAGCATGCAAGTGCATGG 3338
Db ATCTGGCTTCTACGCTGATCCCAATGAGGGAATGTGAGCATGCAAGTGCATGG 231

3339 GCAGCGCTCTCTGTGCAACCAACACAGGCAAGTGTCTTGCAACCAAGGCGCTCAA 3398
Db GCAGCGCTCTCTGTGCAACCAACACAGGCAAGTGTCTTGCAACCAAGGCGCTCAA 291

3399 GGGGACGAGTGCAGCTATGTGAGTAGAATTCGATACCAAGGAACCTCTCAGAGG 3458
Db GGGGACGAGTGCAGCTATGTGAGTAGAATTCGATACCAAGGAACCTCTCAGAGG 351

3459 AACATGTTATATATCTCTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAAGA 3518
Db AACATGTTATATATCTCTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAAGA 411

3519 TGATCGCTATTAACAGCATCAATTTTGTGGCTACTCTCTGACGACAAACAGGATTT 3578
Db TGATCGCTATTAACAGCATCAATTTTGTGGCTACTCTCTGACGACAAACAGGATTT 471

3579 GGACATGTTATCAATGCTCCCAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTT 3638
Db GGACATGTTATCAATGCTCCCAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTT 531

3639 CTCAGCTGGAACCCAGGCTGAGGAAGAGATGCTGTGTTTCAAAAACCAACATTAAAGGA 3698
Db CTCAGCTGGAACCCAGGCTGAGGAAGAGATGCTGTGTTTCAAAAACCAACATTAAAGGA 591

3699 GTACAGAGATGTTCTCTAATGAGAGTTTGATTTTCGCAACCCACCAATATCACTTT 3758
Db GTACAGAGATGTTCTCTAATGAGAGTTTGATTTTCGCAACCCACCAATATCACTTT 651

3759 CTTTGTATTATGCAGTAA 3776
Db CTTTGTATTATGCAGTAA 669

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 3 row: A column: 12
High quality sequence stop: 486.
Location/Qualifiers
1. 486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSSNU484s1-3-A12"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10P"
/clone_lib="SSSNU484s1"
/note="Organ: Stomach; Vector: pTZ18p1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P' by electroporation method.
After analyzing and sequencing about 2,000 - 3,000
colonies in original cDNA library, the abundant cDNAs were
selected and amplified by PCR reaction using vector region
primer including T7 promoter as 5' primer and Nid114 as
3' primer. The PCR products were used as template for
synthesis of biotinylated single stranded RNA by in vitro
transcription reaction. The synthesized RNA probes were
hybridized with antisense single stranded cDNAs prepared
from original library and incubated with avidin-gel.
After removing DNA-RNA hybrids by centrifuge, the
subtracted cDNA libraries were constructed by
transformation of the remaining DNA into competent cells E.
coli Top10P' with electroporation method."

ORIGIN

Query Match 11.3%; Score 485; DB 12; Length 486;
Best Local Similarity 100.0%; Pred. No. 2.5e-220;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATAATGCACTCATCTCAGAGCATGCTCCAGCTCACCTTAACCCCATGGGTC 2571
Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGCTCCAGCTCACCTTAACCCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTCTCTACTGCTGGGAGAGATATGCTCCCATTTACAAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGCTGGGAGAGATATGCTCCCATTTACAAAT 120

QY 2632 AGTTTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGATTTCTGGAAATTTATCA 2691
Db 121 AGTTTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGATTTCTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTGTC 2751
Db 181 GAACCCAGTACTCGGGGACTGAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTGTC 240

QY 2752 TGTGAAGGCGCTGCAAAACACAGTGTCAAGCAGTGCAGGACACCATGTCCTCTGAGGACA 2811
Db 241 TGTGAAGGCGCTGCAAAACACAGTGTCAAGCAGTGCAGGACACCATGTCCTCTGAGGACA 300

QY 2812 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTCAGCAACATGAAGCAG 2871
Db 301 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTCAGCAACATGAAGCAG 360
QY 2872 TGTGTGACTCAATGCGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGACTCAATGCGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
QY 2932 ACGATGAGCACTGCCGCCCTGAAATTTGTTTCAGGCTACTGTACTCTGATTCATTTGCTTG 2991
Db 421 ACGATGAGCACTGCCGCCCTGAAATTTGTTTCAGGCTACTGTACTCTGATTCATTTGCTTG 480
QY 2992 GAGCA 2996
Db 481 GAGCA 485

RESULT 18
BI850289 591 bp mRNA linear EST 10-OCT-2001
LOCUS imagec11.2000/sly246bdr81.v1 NIH_MGC_56 Homo sapiens cDNA clone
DEFINITION IMAGE:4284387 5', mRNA sequence.

ACCESSION BI850289
VERSION BI850289.1 GI:16003776
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and Prange, C.K.

TITLE The I.M.A.G.E. Consortium quality control effort: clone
JOURNAL resequencing for verification
COMMENT Unpublished (2001)
Other ESTs: BF700035
Contact: Prange CK

The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov

This read has been verified (found to hit its original self in the correct orientation), as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.

Plate: L1CM117 row: j column: 4
Seq primer: -21n13
High quality sequence stop: 591.

FEATURES
source
1..591
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4284387"
/issue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccatcgcc); Site_2: SfiI (ggccatcgcc); SfiI (ggccatcgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dt(30)BN-3' (where B = A, C, G or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 11.0%; Score 471; DB 12; Length 591;
Best Local Similarity 99.8%; Pred. No. 1.3e-213;
Matches 591; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2636 TACTACAGTGGATGCGCTCTGAGCCAGTGTGCTGGATTTCTGTGAAATTTATCAGAAC 2695
Db 1 TACTACAGTGGATGCGCTCTGAGCCAGTGTGCTGGATTTCTGTGAAATTTATCAGAAC 60
QY 2696 CCAGTACTCGGGGACTGAGGCTGGAACCTGATCAACCCACTCAATGCTAGTGTCTGTG 2755
Db 61 CCAGTACTCGGGGACTGAGGCTGGAACCTGATCAACCCACTCAATGCTAGTGTCTGTG 120
QY 2756 AAAGGCTTCAAAACCAACAGTGTCTTAAGCAGTCCGGACACCATGTGCTTTGAGGACAGCAT 2815
Db 121 AAAGGCTTCAAAACCAACAGTGTCTTAAGCAGTCCGGACACCATGTGCTTTGAGGACAGCAT 179
QY 2816 GTGGAGATTGCCACCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2875
Db 180 GTGGAGATTGCCACCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 239
QY 2876 TGGACTTCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAAATGATATACGA 2935
Db 240 TGGACTTCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAAATGATATACGA 299
QY 2936 TGAGCACTGCCCCCTGAAATTTGTTTACGGCTACTGTACCTGTAGTCTATTCCTTGGAGC 2995
Db 300 TGAGCACTGCCCCCTGAAATTTGTTTACGGCTACTGTACCTGTAGTCTATTCCTTGGAGC 359
QY 2996 AACCGGCTGTGGCTGTACTGTATCCAGCAATCTGGCAAGGAAATGCATAGAGG 3055
Db 360 AACCGGCTGTGGCTGTACTGTATCCAGCAATCTGGCAAGGAAATGCATAGAGG 419
QY 3056 GTTCTTATAAAGGACAGTGAAGATGCTTCCCAAGCCCTACAGAAATTTCTATCCAC 3115
Db 420 GTTCTTATAAAGGACAGTGAAGATGCTTCCCAAGCCCTACAGAAATTTCTATCCAC 479
QY 3116 AGCCCTGTCTCAATTCACGATGTGCTAGAGGACAGAGATACACTGGTCTTTCATTC 3175
Db 480 AGCCCTGTCTCAATTCACGATGTGCTAGAGGACAGAGATACACTGGTCTTTCATTC 539
QY 3176 ACTGTCCAGCTTGGCAATCAACGGCCACAGTAAATGATCAATCAGAGCAT 3227
Db 540 ACTGTCCAGCTTGGCAATCAACGGCCACAGTAAATGATCAATCAGAGCAT 591

RESULT 19

BF853006

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Shogun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Shogun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Shogun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Shogun sequencing of the human transcriptome with ORF expressed

Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR&st2=MR2-EN0093-191200-001-b07&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 598.
Location/Qualifiers
1. 598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES

source

ORIGIN

Query Match 10.1%; Score 435; DB 10; Length 598;
Best Local Similarity 99.8%; Pred. No. 2.3e-196;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3282 TGGCTTCTACGGTATCCCAATGAGGGAATGTCAGCCATGCAAGTGCATATGGCA 3341
DB 5 TGGCTTCTACGGTATCCCAATGAGGGAATGTCAGCCATGCAAGTGCATATGGCA 64
QY 3342 CGCGTCTCTGCAACACACACGCGGCAAGTGTCTTGCAACACACGCGGCGTCAAGG 3401
DB 65 CGCGTCTCTGCAACACACACGCGGCAAGTGTCTTGCAACACACGCGGCGTCAAGG 124
QY 3402 GGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAAC 3461
DB 125 GGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAAC 184
QY 3462 ATGTTATTATCTCTTATTGACTATCAGTTCACTTGTAGTCTATCCAGGAGTGA 3521
DB 185 ATGTTATTATCTCTTATTGACTATCAGTTCACTTGTAGTCTATCCAGGAGTGA 244
QY 3522 TCGCTATTACACAGCTATCAATTTTGGCTACTCTGACCAACCAACAGGGATTGGA 3581
DB 245 TCGCTATTACACAGCTATCAATTTTGGCTACTCTGACCAACCAACAGGGATTGGA 304
QY 3582 CATGTTATCAATGCTCCAGAAATTTCAACTCAACATCCTGGGCTGCCAGTTTCTC 3641
DB 305 CATGTTATCAATGCTCCAGAAATTTCAACTCAACATCCTGGGCTGCCAGTTTCTC 364
QY 3642 AGCTGGACCCAGGCTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAGGAGTA 3701
DB 365 AGCTGGACCCAGGCTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAGGAGTA 424
QY 3702 CAAGATAGTTTCTCTAATAGAAAGTTTGAATTTTGCACACCCCAATATCACTTTCTT 3761
DB 425 CAAGATAGTTTCTCTAATAGAAAGTTTGAATTTTGCACACCCCAATATCACTTTCTT 484
QY 3762 TGTTTA 3767
DB 485 TGTTTA 490

RESULT 20

CB242644/C

LOCUS

CB242644

DEFINITION

UI-CF-FNO-afs-i-02-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone

ACCESSION

CB242644

VERSION CB242644.1 GI:28364288
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 746)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE 97044477
PUBMED 8889548
COMMENT
Contact: McCray, PB
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 16-237, >L1PA2#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. 746
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="UI-CF-FNO-afs-i-02-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pTV73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUL) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bentso-soares@uiowa.edu
TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h
TAG LIB=UI-CF-FNO
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 9.9%; Score 426; DB 14; Length 746;
Best Local Similarity 100.0%; Pred. No. 4.9e-192;
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3376 TTCTGCACCAACAGGGCGTCAAGGGGACGAGTCCAGCTATGAGGTAGAAATCGA 3435
DB 661 TTCTGCACCAACAGGGCGTCAAGGGGACGAGTCCAGCTATGAGGTAGAAATCGA 602
QY 3436 TACCAGGAACCCCTCTCAGAGGAATGATCGCTATTACACAGTATCAATTTTGTGGTACT 3495
DB 601 TACCAGGAACCCCTCTCAGAGGAATGATCGCTATTACACAGTATCAATTTTGTGGTACT 542
QY 3496 ACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGTATCAATTTTGTGGTACT 3555
DB 541 ACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGTATCAATTTTGTGGTACT 482
QY 3556 CCTGACGACAAACAGGGATTGGACATGTTTCATCAATGCTCCCAAGATTTCAACCTC 3615
DB 481 CCTGACGACAAACAGGGATTGGACATGTTTCATCAATGCTCCCAAGATTTCAACCTC 422


```

3616 AACATCACTGGGTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTGTT 3675
|||||
421 AACATCACTGGGTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTGTT 362
|||||
3676 GTTTCAAAACCAACATAGGAGTACAAAGATAGTTTCTCTATGAGAGATTTGATTTT 3735
|||||
361 GTTTCAAAACCAACATAGGAGTACAAAGATAGTTTCTCTATGAGAGATTTGATTTT 302
|||||
3736 CGCAACCAACCAATATACATCTTTCTTTTATGTCAGTAAATTTCACTGGCCCATCAA 3795
|||||
301 CGCAACCAACCAATATACATCTTTCTTTTATGTCAGTAAATTTCACTGGCCCATCAA 242
|||||
3796 ATTCAG 3801
|||||
241 ATTCAG 236
|||||

RESULT 21
BF853681
LOCUS BF853681 595 bp mRNA linear EST 16-JAN-2001
DEFINITION MR2-EN0093-211200-003-a05 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853681
VERSION BF853681.1 GI:12241425
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 595)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
211200-003-a05&t3=2000-12-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.
FEATURES
Location/Qualifiers
1..595
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 9.9%; Score 425; DB 10; Length 595;
Best Local Similarity 99.8%; Pred. No. 1.4e-151;

```

```

Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3282 TGGCTTTACGGTGTATCCCAACCAATGGAGGAAATGTTCAGCCATGCAATGGGCA 3341
DB 6 TGGCTTTACGGTGTATCCCAACCAATGGAGGAAATGTTCAGCCATGCAATGGGCA 65
QY 3342 CCGCTCTCTGTGCAACCAACACCGGCAAGTGTCTTCACCACCAAGGGGCTCAAGG 3401
DB 66 CCGCTCTCTGTGCAACCAACACCGGCAAGTGTCTTCACCACCAAGGGGCTCAAGG 125
QY 3402 CGACGAGTGGCCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAC 3461
DB 126 GACGAGTGGCCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAC 185
QY 3462 ATGTTATTATATCTCTCTTATTGACTATCAGTTACCTTTAGTCTATCCAGGAAGATGA 3521
DB 186 ATGTTATTATATCTCTCTTATTGACTATCAGTTACCTTTAGTCTATCCAGGAAGATGA 245
QY 3522 TCGCTATTACACAGCTATCAATTTTGTGCTACTCTCGACGAACAAACACAGGATTTGA 3581
DB 246 TCGCTATTACACAGCTATCAATTTTGTGCTACTCTCGACGAACAAACACAGGATTTGA 305
QY 3582 CATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTC 3641
DB 306 CATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTC 365
QY 3642 AGCTGGACCCAGGCTGGAGAGAGATGCTGTTTTCATTTTCATTTTCATTTTCATTTT 3701
DB 366 AGCTGGACCCAGGCTGGAGAGAGATGCTGTTTTCATTTTCATTTTCATTTTCATTTT 425
QY 3702 CAAAGATAGTTTCTCTAATGAGAAGTTTTCATTTTCGCAACCAACCAATATACATT 3757
DB 426 CAAAGATAGTTTCTCTAATGAGAAGTTTTCATTTTCGCAACCAACCAATATACATT 481

RESULT 22
BF183626/c
LOCUS BF183626 744 bp mRNA linear EST 15-JUL-2003
DEFINITION UI-R-EU0-azn-a-22-0-UI-s1 NCI-CGAP_Carl Homo sapiens cDNA clone
IMAGE:5850909 3', mRNA sequence.
ACCESSION BF183626
VERSION BF183626.1 GI:20359183
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 744)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 16-237, >h1P2#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
Location/Qualifiers
1..744
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5850909"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"

```



```

QY 3766 TATGTCAGTAATTTCACTGGCCC 3789
Db 87 TATGTCAGTAATTTCACTGGCCC 64

RESULT 24
BF853962
LOCUS BF853962 544 bp mRNA linear EST 16-JAN-2001
DEFINITION MR2-EN0093-261200-004-g08 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853962
VERSION BF853962.1 GI:12241706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 544)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
261200-004-g08&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 544.
Location/Qualifiers
1..544
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

FEATURES
source
Query Match 9.3%; Score 400; DB 10; Length 544;
Best Local Similarity 99.8%; Pred. No. 1.3e-179;
Matches 450; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3336 TGGGACGCGTCTGTGACACACACAGGGGCAAGTGTCTTGCACCAACCAAGGCGT 3395
Db 63 TGGGACGCGTCTGTGACACACACAGGGGCAAGTGTCTTGCACCAACCAAGGCGT 122

QY 3396 CAAGGGGACAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGAAACCTCTTCAG 3455
Db 123 CAAGGGGACAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGAAACCTCTTCAG 182

QY 3456 AGGAACATGTTATTACTCTTCTATTGACTATCACTTACCTTTAGTCTATCCACGGA 3515
Db 183 AGGAACATGTTATTACTCTTCTATTGACTATCACTTACCTTTAGTCTATCCACGGA 242

```

```

QY 3516 AGATGATCGCTATTACACAGCTATCAATTTTGGTACTCTCTGAGCAACAAACAGGGA 3575
Db 243 AGATGATCGCTATTACACAGCTATCAATTTTGGTACTCTCTGAGCAACAAACAGGGA 302

QY 3576 TTTCGACATGTTTCATCAATGCTCCCAAGAAATTCACCTCAACATCACCTGGCTGCCAG 3635
Db 303 TTTCGACATGTTTCATCAATGCTCCCAAGAAATTCACCTCAACATCACCTGGCTGCCAG 362

QY 3636 TTTCCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTA 3695
Db 363 TTTCCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTA 422

QY 3696 GGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTCAATTTTCGCAACCCCAATATCAC 3755
Db 423 GGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTCAATTTTCGCAACCCCAATATCAC 482

QY 3756 TTTCCTTTGTTTATGTCAGTAATTTCACTGG 3786
Db 483 TTTCCTTTGTTTATGTCAGTAATTTCACTGG 513

```

```

RESULT 25
CA306924/c
LOCUS CA306924 643 bp mRNA linear EST 01-NOV-2002
DEFINITION UI-H-Ft1-bhu-e-21-0-UI-s1 NCI CGAP Ftl Homo sapiens cDNA clone
UI-H-Ft1-bhu-e-21-0-UI 3', mRNA sequence.
ACCESSION CA306924
VERSION CA306924.1 GI:24469978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 643)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bentso-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 16-237 >L1PA2#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..643
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Ftl"
/tissue_type="Aveolar Macrophage"
/lab_host="DH10B (Life Technologies)"
/dev_stage="Adult"

```

```

FEATURES
source
1..643
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP Ftl is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag

```

sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCCATGCCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
TAG_R1SUB=Human Lung Alveolar Macrophage
TAG_LIB=UI-H-FT1
TAG_SEQ=GGCCATGCCG"

ORIGIN

Query Match 9.3%; Score 399; DB 14; Length 643;
Best Local Similarity 100.0%; Pred. No. 4e-179;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3403 GACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACA 3462
Db 634 GACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACA 575
QY 3463 TGTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGAT 3522
Db 574 TGTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGAT 515
QY 3523 CGCTATTACACAGCTATCAATTTTGGCTACTCTCTGACGAACAAAACAGGGATTTGGAC 3582
Db 514 CGCTATTACACAGCTATCAATTTTGGCTACTCTCTGACGAACAAAACAGGGATTTGGAC 455
QY 3583 ATGTTTCATCAATGCTCCAGAAATTCACCTCAACATCAGCTGGGCTGCAGTTTCTCA 3642
Db 454 ATGTTTCATCAATGCTCCAGAAATTCACCTCAACATCAGCTGGGCTGCAGTTTCTCA 395
QY 3643 GCTGGAACCCAGGCTGGAGAGAGATGCTGCTGTTTCAAAAACCAACATTAAGGAGTAC 3702
Db 394 GCTGGAACCCAGGCTGGAGAGAGATGCTGCTGTTTCAAAAACCAACATTAAGGAGTAC 335
QY 3703 AAAGATAGTTTCTCTAATGAGAAGTTGATTTTCCAAACCAACCAATTAATCATTCTTT 3762
Db 334 AAAGATAGTTTCTCTAATGAGAAGTTGATTTTCCAAACCAACCAATTAATCATTCTTT 275
QY 3763 GTTTATGTCAGTATTTTCACTGACCTGCCATCAAAATTCAG 3801
Db 274 GTTTATGTCAGTATTTTCACTGACCTGCCATCAAAATTCAG 236

RESULT 26
AY418588 3671 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY418588
VERSION AY418588.1 GI:39774548
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
1 (bases 1 to 3671)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3671)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
JOURNAL
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
source
1..3671
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>3671
/gene="ATRN"
/locus_tag="HCM6607"

ORIGIN

Query Match 9.1%; Score 390; DB 29; Length 3671;
Best Local Similarity 99.1%; Pred. No. 1.2e-174;
Matches 690; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3470 ATACTCTTCTTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAGATGATGCTATT 3529
Db 2976 ATACTCTTCTTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAGATGATGCTATT 3035
QY 3530 ACACAGCTATCAATTTTGGCTACTCTGACGAACAAAACAGGATTTGGACATGTTC 3589
Db 3036 ACACAGCTATCAATTTTGGCTACTCTGACGAACAAAACAGGATTTGGACATGTTC 3095
QY 3590 TCAATGCTCTCCAAAGAAATTTCAACCTCAACATCACTGGCTGCCAGTTTTCAGCTGAA 3649
Db 3096 TCAATGCTCTCCAAAGAAATTTCAACCTCAACATCACTGGCTGCCAGTTTTCAGCTGAA 3155
QY 3650 CCCAGCTGGAGAGAGATGCTGTTTTCAAAACCAACATTAAGAGATACAAAGTA 3709
Db 3156 CCCAGCTGGAGAGAGATGCTGTTTTCAAAACCAACATTAAGAGATACAAAGTA 3215
QY 3710 GTTCTCTTAATCAGAAAGTTTGAATTTTCGCAACCAACCAATTAATGTTTATG 3769
Db 3216 GTTCTCTTAATCAGAAAGTTTGAATTTTCGCAACCAACCAATTAATGTTTATG 3275
QY 3770 TCAGTAATTTTCACTGGCCCATCAAAATTCAGATTCAGTTTCTCTCAGCACAGCAATTTTA 3829
Db 3276 TCAGTAATTTTCACTGGCCCATCAAAATTCAGATTCAGTTTCTCTCAGCACAGCAATTTTA 3335
QY 3830 TGGACCTGGTACAGTTCTTCGTGACTTTTTCAGTTGTTTCTCTCTCTCTCTCTCTCTCT 3899
Db 3336 TGGACCTGGTACAGTTCTTCGTGACTTTTTCAGTTGTTTCTCTCTCTCTCTCTCTCTCT 3395
QY 3890 CTGCTGTGTTTGGAGATCAAAACAAAGTTTGTGGGCTCCAGACGTAGAGACAACTTC 3949
Db 3396 CTGCTGTGTTTGGAGATCAAAACAAAGTTTGTGGGCTCCAGACGTAGAGACAACTTC 3455
QY 3950 TTTCAGAGATGCAACAGATGGCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4009
Db 3456 TTTCAGAGATGCAACAGATGGCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3515
QY 4010 CAGATGAGGAGCT 4069
Db 3516 CAGATGAGGAGCT 3575
QY 4070 CACTGGAGCCGTGTTTTCGCAACAAAGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4129
Db 3576 CACTGGAGCCGTGTTTTCGCAACAAAGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3635
QY 4130 GAGGCTGGTGGCATCCCT 4165
Db 3636 GAGGCTGGTGGCATCCCT

RESULT 27
CA502816/c
LOCUS
DEFINITION UI-CF-FNO-afg-d-10-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
ACCESSION CA502816
VERSION CA502816.1 GI:24993770
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

CA502816 739 bp mRNA linear EST 14-NOV-2002
UI-CF-FNO-afg-d-10-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
UI-CF-FNO-afg-d-10-0-UI 3', mRNA sequence.
CA502816.1 GI:24993770
EST.
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 739)
Bonald,M.F., Lemmon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@iowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 16-237, >LPA2HLINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1..739
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-afg-d-10-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUI). The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h
TAG_LIB=UI-CF-FNO
TAG_SEQ=CTGCTCAGGT"

ORIGIN
Query Match 9.1%; Score 389; DB 14; Length 739;
Best Local Similarity 100.0%; Pred. No. 2.5e-174;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3413 AGCTATGTGAGGTAGAAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATTATA 3472
DB 624 AGCTATGTGAGGTAGAAAATCGATACCAAGAAACCTCTCAGAGGACATGTTATTATA 565
QY 3473 CTCTTCTATTGACTATCAGTTACCTTTAGCTTATGCCAGGAAGATGCTGTTATACA 3532
DB 564 CTCTTCTATTGACTATCAGTTACCTTTAGCTTATGCCAGGAAGATGCTGCTATTACA 505
QY 3533 CAGCTATCAATTTTGTGGCTACTCTCAGCAACAAACAGGATTTGGACATGTTCTATCA 3592
DB 504 CAGCTATCAATTTTGTGGCTACTCTCAGCAACAAACAGGATTTGGACATGTTCTATCA 445
QY 3593 ATGCCTCCAGAAATTTCAACCTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAACCC 3652
DB 444 ATGCCTCCAGAAATTTCAACCTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAACCC 385
QY 3653 AGGCTGAGAGAGATCGCTGTGTTTCAAAAACCAACATTAGGAGTACAAAGATGTT 3712
DB 384 AGGCTGAGAGAGATCGCTGTGTTTCAAAAACCAACATTAGGAGTACAAAGATGTT 325

QY 3713 TCTCTAATGAGAAGTTTGTATTTTGGCAACCAACCAAAATATACATTTCTTTTGTATGCA 3772
DB 324 TCTCTAATGAGAAGTTTGTATTTTGGCAACCAACCAAAATATACATTTCTTTTGTATGCA 265
QY 3773 GTAAATTTCACTGGCCCATCAAAATTCAG 3801
DB 264 GTAAATTTCACTGGCCCATCAAAATTCAG 236

RESULT 28
BG563729
LOCUS 602584512F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4712114 5', mRNA sequence.
DEFINITION BG563729
ACCESSION BG563729
VERSION BG563729.1 GI:13571381
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 674)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Ph.D.
TITLE National Institutes of Health, Ph.D.
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1553 row: h column: 03
High quality sequence stop: 674.
Location/Qualifiers
FEATURES
source
1..674
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4712114"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_76"
/notes="Organ: Liver; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccattggcc); Site 2: SfiI (ggccattggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCATATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 8.9%; Score 382; DB 12; Length 674;
Best Local Similarity 99.6%; Pred. No. 5.6e-171;
Matches 552; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 2568 GGTGCGCCTTCGGAAGATCAATGTCTCTACTGGTGGGAAGATATGCCCATTTTAC 2627
DB 1 GGTGCGCCTTCGGAAGATCAATGTCTCTACTGGTGGGAAGATATGCCCATTTTAC 60
QY 2628 AAATAGTTTACTACAGTGGATGCCCTCTGAGCCAGTATGCTGGAATTTT 2687
DB 61 AAATAGTTTACTACAGTGGATGCCCTCTGAGCCAGTATGCTGGAATTTT 120
QY 2688 ATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCACTCAACCCACTCAATGGTAG 2747
DB 121 ATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCACTCAACCCACTCAATGGTAG 180

ORIGIN

Average insert size 1.1 kb. Library constructed by Life Technologies."

Query Match 8.5%; Score 363; DB 10; Length 991;
Best Local Similarity 99.8%; Pred. No. 7.6e-162;
Matches 533; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1536 GGCCCTATACGTTTCATGGTGGCTTACAAAGGCTTTTCAGTGGCAATPAAGTACCGGCTTCGAGA 1595
DB 38 GGCCCTATACGTTTCATGGTGGCTTACAAAGGCTTTTCAGTGGCAATPAAGTACCGGCTTCGAGA 97

QY 1596 TGATCTCTACCGATATAGTGTGATACCCAGATGTCAGCAATCTTAAAGACAGCGGAT 1655
DB 98 TGATCTCTACCGATATAGTGTGATACCCAGATGTCAGCAATCTTAAAGACAGCGGAT 157

QY 1656 TTTCCGTTACTTGCACACAGCTGTGATGAGTGAACCATCTCTGTTGGGGGAAA 1715
DB 158 TTTCCGTTACTTGCACACAGCTGTGATGAGTGAACCATCTCTGTTGGGGGAAA 217

QY 1716 CACACAAATGACATCTATAGCCATGGCGGCAATGTTCTCTTCAGATTTCATGGC 1775
DB 218 CACACAAATGACATCTATAGCCATGGCGGCAATGTTCTCTTCAGATTTCATGGC 277

QY 1776 CTATGACATTCGCTGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGT 1835
DB 278 CTATGACATTCGCTGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGT 337

QY 1836 CAACAGATTTGGCGATTCAGACGCTTTACACACAGCACCATGATGTGTTGCGTGGTTT 1895
DB 338 CAACAGATTTGGCGATTCAGACGCTTTACACACAGCACCATGATGTGTTGCGTGGTTT 397

QY 1896 CAATAGTCTCCTCTCAGACGACATCTGTTGATTCACCTCGGACAGCTGTG-ATGCCATC 1954
DB 398 CAATAGTCTCCTCTCAGACGACATCTGTTGATTCACCTCGGACAGCTGTGATTCGCTGATC 457

QY 1955 GGAGTGAAGCGCTTGTGTAGCAGGACCTGATTCGGTGTGTGTGGAACACAGGTT 2014
DB 458 GGAGTGAAGCGCTTGTGTAGCAGGACCTGATTCGGTGTGTGTGGAACACAGGTT 517

QY 2015 CGTCTAGTGTATCTCGTGGCGCTGCAATGATGACAGAGAAAAGTTAAA 2069
DB 518 CGTCTAGTGTATCTCGTGGCGCTGCAATGATGACAGAGAAAAGTTAAA 572

RESULT 31
BF963698/c
LOCUS
DEFINITION PM4-NN1204-221200-001-903 NN1204 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF963698
VERSION BF963698.1 GI:12380973
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 484)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., Gellivelira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBMED
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-NN1204-221200-001-903&t3=2000-12-22&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 350.
Location/Qualifiers
1. 484
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/dev_stage="Adult"
/clone_lib="NN1204"
/note="Organ: nervous normal; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

ORIGIN

Query Match 8.4%; Score 360; DB 10; Length 484;
Best Local Similarity 99.8%; Pred. No. 1.8e-160;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 3130
DB 484 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 425

QY 3131 CCAGCATGTGCTAGAGCAGCAGATACAACTGGTCTTTCATTCACATGTCAGCTGCC 3190
DB 424 CCAGCATGTGCTAGAGCAGCAGATACAACTGGTCTTTCATTCACATGTCAGCTGCC 365

QY 3191 AATGCAAGGGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 3250
DB 364 AATGCAAGGGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 305

QY 3251 CCACAGGCAAGCAGTGGCAGACCTGCATATCTGGCTTCTACGGTATCCACCAATGAG 3310
DB 304 CCACAGGCAAGCAGTGGCAGACCTGCATATCTGGCTTCTACGGTATCCACCAATGAG 245

QY 3311 GGAATGTCAGCATGCAAGTGAAGTGGCAGCGCTCTGTGCAACACCAACCGGCA 3370
DB 244 GGAATGTCAGCATGCAAGTGAAGTGGCAGCGCTCTGTGCAACACCAACCGGCA 185

QY 3371 ACTGCTTTCTGCACCAAGGGCGTCAAGGGGGACGAGTGCAGCTATGTGAGGTAGAAA 3430
DB 184 ACTGCTTTCTGCACCAAGGGCGTCAAGGGGGACGAGTGCAGCTATGTGAGGTAGAAA 125

QY 3431 ATCGATACCAAGAAACCTCTCAGAGGACATGTTATATCTTCTTA 3481
DB 124 ATCGATACCAAGAAACCTCTCAGAGGACATGTTATATCTTCTTA 74

RESULT 32
BF313142
LOCUS
DEFINITION 601897661F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4126959 5', mRNA sequence.
ACCESSION BF313142
VERSION BF313142.1 GI:11261090
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 953)
NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLM1017 row: j column: 16
 High quality sequence stop: 612.
FEATURES Location/Qualifiers
 source
 1..953
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4126959"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_19"
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN
 Query Match 8.3%; Score 355; DB 10; Length 953;
 Best Local Similarity 99.6%; Pred. No. 5.1e-158;
 Matches 525; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1180 GGTCAATCTTTGGCATTATACAGAGTAAATTTACATGATGGAGAAAATGATCCA 1239
 Db 1 GGTCAATCTTTGGCATTATACAGAGTAAATTTACATGATGGAGAAAATGATCCA 60
 QY 1240 ACTGGGAATGTGACCAATGATGTCAGAGTTTTCACATTCATATGATGTCGGTGTG 1299
 Db 61 ACTGGGAATGTGACCAATGATGTCAGAGTTTTCACATTCATATGATGTCGGTGTG 120
 QY 1300 TTGACCCCTTAAGGCAAGGAGCAGTATGTCAGTGGTGGGCACTCTGCACACATTTGTACA 1359
 Db 121 TTGACCCCTTAAGGCAAGGAGCAGTATGTCAGTGGTGGGCACTCTGCACACATTTGTACA 180
 QY 1360 CTGAGGAATGCCGAGTGGTCATCTTTGGTCACTGGCCCTCTCTATGGATAT 1419
 Db 181 CTGAGGAATGCCGAGTGGTCATCTTTGGTCACTGGCCCTCTCTATGGATAT 240
 QY 1420 ATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACCCAG 1479
 Db 241 ATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACCCAG 300
 QY 1480 GGTGCCCTTTGCAAGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCGGCC 1539
 Db 301 GGTGCCCTTTGCAAGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCA-GGCC 359
 QY 1540 CTATACCTTCATGGTGGCTACAGGCTTTTCAGTGGCAATAGTACCGGCTTTCAGATGAT 1599
 Db 360 CTATACCTTCATGGTGGCTACAGGCTTTTCAGTGGCAATAGTACCGGCTTTCAGATGAT 419
 QY 1600 CTCACCGATATGATGTGGATACCCAGATGTGGACCATCTTTAAGACAGCCGATTTTC 1659
 Db 420 CTCACCGATATGATGTGGATACCCAGATGTGGACCATCTTTAAGACAGCCGATTTTC 479
 QY 1660 CGTTACTTGCACACAGCTGTGATGATGAGTGGGAACCATGCTGGTGT 1706
 Db 480 CGTTACTTGCACACAGCTGTGATGATGAGTGGGAACCATGCTGGTGT 526

RESULT 33

BF853965
 LOCUS MR2-EN0093-261200-004-h12 EN0093 Homo sapiens cDNA, mRNA linear EST 16-JAN-2001
 DEFINITION BP853965
 ACCESSION BP853965
 VERSION BP853965.1 GI:12241709
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 560)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR2&t2=MR2-EN0093-261200-004-h12&t3=2000-12-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 559.
FEATURES Location/Qualifiers
 1..560
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="EN0093"
 /note="Organ: lung normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196, 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN
 Query Match 8.2%; Score 353; DB 10; Length 560;
 Best Local Similarity 99.8%; Pred. No. 4.1e-157;
 Matches 403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3337 GGGCAGCGTCTCTGTGCAACACCAACGCGGCAAGTCTTTCACACCAAGGGGTC 3396
 Db 62 GGGCAGCGTCTCTGTGCAACACCAACGCGGCAAGTCTTTCACACCAAGGGGTC 121
 QY 3397 AAGGGGACGAGTGCAGCTATGTAGTGAATAATCGATACCAAGGAACCCCTCTCAGA 3456
 Db 122 AAGGGGACGAGTGCAGCTATGTAGTGAATAATCGATACCAAGGAACCCCTCTCAGA 181
 QY 3457 GGAACATGTTATTATCTCTTCTTTATTGACTATACAGTTACCTTTAGTCTATCCAGAA 3516
 Db 182 GGAACATGTTATTATCTCTTCTTTATTGACTATACAGTTACCTTTAGTCTATCCAGAA 241
 QY 3517 GATGATCGCTATTACAGCTATCAATTTTGTGCTACTCTGACGAACAAACAGGAT 3576
 Db 242 GATGATCGCTATTACAGCTATCAATTTTGTGCTACTCTGACGAACAAACAGGAT 301
 QY 3577 TTGGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGTGCCAGT 3636

```

Db      302 TTGACAGTGTTCATCAATGCTCCAGAAATTCACATCTCACATCACTGGGCTGCCAT 361
Qy      3637 TTCTAGCTGGAACCCAGGCTGGAGAGAGATCCCTGTGTTTCAAAAACCAACATTAAG 3696
Db      362 TTCTAGCTGGAACCCAGGCTGGAGAGAGATCCCTGTGTTTCAAAAACCAACATTAAG 421
Qy      3697 GAGTACAAAGATAGTTCTCTAATGAGAGCTTTGATTTCGCAA 3740
Db      422 GAGTACAAAGATAGTTCTCTAATGAGAGCTTTGATTTCGCAA 465

RESULT 34
LOCUS   BG673763
DEFINITION 9010 ciliated epithelial cDNA cell library Homo sapiens cDNA 5',
mRNA sequence.
ACCESSION BG673763
VERSION   BG673763.2 GI:18490081
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 390)
AUTHORS Maiti,A.K., Jorissen,M. and Bouvagnet,P.
TITLE Isolation, in silico characterization and chromosomal localization
of a group of cDNAs from ciliated epithelial cells after in vitro
ciliogenesis
JOURNAL Genome Biol. 2 (7), RESEARCH0026 (2001)
MEDLINE 21407920
PUBMED 11516339
COMMENT On May 1, 2001 this sequence, version replaced gi:13919402.
Contact: Maiti AK
Laboratory de Genetic Moleculaire Humaine, Faculty de Pharmacy
University Cl. Bernard, F69373 Lyon cedex, France
8 Avenue Rockefeller, F69373 Lyon cedex, France
Tel: (33) 478 77 44 25
Fax: (33) 478 77 75 68
Email: amit.maiti@medecine.unige.ch
Seq primer: M13 Forward and reverse.
Location/Qualifiers
1..390
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Nasal biopsy"
/cell_type="cultured ciliated epithelial cells after in
vitro Ciliogenesis"
/dev_stage="Adult"
/clone_lib="ciliated epithelial cDNA cell library"
/note="vector: PCR2.1; cloned RT-PCR fragment from
isolated total RNA"

FEATURES
source
1..390
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Nasal biopsy"
/cell_type="cultured ciliated epithelial cells after in
vitro Ciliogenesis"
/dev_stage="Adult"
/clone_lib="ciliated epithelial cDNA cell library"
/note="vector: PCR2.1; cloned RT-PCR fragment from
isolated total RNA"

ORIGIN
Query Match 8.2%; Score 352; DB 12; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.le-156;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      852 TCCTCACTGTACAGACAACCTGTGTTTTCCTCATCGAGGCATCTGCAATTCAGTGTATGT 911
Db      1 TCCTCACTGTACAGACAACCTGTGTTTTCCTCATCGAGGCATCTGCAATTCAGTGTATGT 60
Qy      912 CAGAGGATGCTTCCTGCTTCTCAGACTGGCAGGTCCTGGATGTTCAGTTCCTGTACCAGC 971
Db      61 CAGAGGATGCTTCCTGCTTCTCAGACTGGCAGGTCCTGGATGTTCAGTTCCTGTACCAGC 120
Qy      972 TAACAGATCATTTTGGACTCGAGAGGAATATTCCTAAAGCTCCCGAGAGCATCTCA 1031
Db      121 TAACAGATCATTTTGGACTCGAGAGGAATATTCCTAAAGCTCCCGAGAGCATCTCA 180
Qy      1032 TAAAGCTGTGTCATATGGAACCATATATGTGGGTGTTGGAGGATATATGTTCAACCACTC 1091

```

```

Db      181 TAAAGCTGTGTCATATGTCGAAACATTTATGTGGGTGTTTGGAGGATATATGTTCAACCACTC 240
Qy      1092 AGATTATTAACATGTTCTTAGCGTATGACCTTCTTAGGAGTGGCTTCCTCACTAAACCG 1151
Db      241 AGATTATTAACATGTTCTTAGCGTATGACCTTCTTAGGAGTGGCTTCCTCACTAAACCG 300
Qy      1152 TTCTGTGAACAATGTGGTGTGTAGATATGTCATCTTTGGCAATATACAAG 1203
Db      301 TTCTGTGAACAATGTGGTGTGTAGATATGTCATCTTTGGCAATATACAAG 352

RESULT 35
LOCUS   BM981492/c
DEFINITION 681 bp mRNA linear EST 21-FEB-2003
UI-CF-EN1-adh-j-10-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-adh-j-10-0-UI 3', mRNA sequence.
ACCESSION BM981492
VERSION   BM981492.1 GI:19604039
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 16-237, >LIP2#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-adh-j-10-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pTT3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an Ecor I adaptor, digested with Not
I, and cloned directionally into pTT3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTGCTCAGGT."

```

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG LIB=UI-CF-ENL
TAG_SEQ=CTGCTCAGGT

ORIGIN

Query Match 8.1%; Score 347; DB 12; Length 681;
Best Local Similarity 99.7%; Pred. No. 3.2e-154;
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3404 ACAGTCCACGCTATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACAT 3463
3404 ACAGTCCACGCTATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACAT 3463
633 ACAGTCCACGCTATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACAT 574
633 ACAGTCCACGCTATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACAT 574
3464 GTTATTATACCTCTCTTATTGACATATCAGTTTACCTTTAGTCTATCCAGGAGATGATC 3523
3464 GTTATTATACCTCTCTTATTGACATATCAGTTTACCTTTAGTCTATCCAGGAGATGATC 3523
573 GTTATTATACCTCTCTTATTGACATATCAGTTTACCTTTAGTCTATCCAGGAGATGATC 514
573 GTTATTATACCTCTCTTATTGACATATCAGTTTACCTTTAGTCTATCCAGGAGATGATC 514
3524 GCTATTACAGCTATCATATTTTGTGCTACTCTCTGACGAAACAAACAGGGATTTGGACA 3583
3524 GCTATTACAGCTATCATATTTTGTGCTACTCTCTGACGAAACAAACAGGGATTTGGACA 3583
513 GCTATTACAGCTATCATATTTTGTGCTACTCTCTGACGAAACAAACAGGGATTTGGACA 454
513 GCTATTACAGCTATCATATTTTGTGCTACTCTCTGACGAAACAAACAGGGATTTGGACA 454
3584 TGTTTCATCAATGCTCCAAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAG 3643
3584 TGTTTCATCAATGCTCCAAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAG 3643
453 TGTTTCATCAATGCTCCAAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAG 394
453 TGTTTCATCAATGCTCCAAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAG 394
3644 CTGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 3703
3644 CTGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 3703
393 CTGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 334
393 CTGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 334
3704 AAGATAGTTTCTTAATGAGAGATTTGATTTTGCACACCCCAATATACATTTCTTTG 3763
3704 AAGATAGTTTCTTAATGAGAGATTTGATTTTGCACACCCCAATATACATTTCTTTG 3763
333 AAGATAGTTTCTTAATGAGAGATTTGATTTTGCACACCCCAATATACATTTCTTTG 274
333 AAGATAGTTTCTTAATGAGAGATTTGATTTTGCACACCCCAATATACATTTCTTTG 274
3764 TTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 3801
3764 TTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 3801
273 TTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 236
273 TTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 236

RESULT 36
BE259981
LOCUS 601148539F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163953 5',
DEFINITION mRNA sequence.
ACCESSION BE259981
VERSION BE259981.1 GI:9130910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 493)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM125 row: m column: 10
High quality sequence stop: 490.
Location/Qualifiers
1. 493
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3163953"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"

FEATURES
source
1. 493
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3163953"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"

/clone lib="NIH_MGC_19"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into ECORI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 7.8%; Score 335; DB 10; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.7e-148;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1354 GTTACACTGAAGAATGGCGAGTGTCTATCTTTGGTCACTGCCCTCTCTAT 1413
1354 GTTACACTGAAGAATGGCGAGTGTCTATCTTTGGTCACTGCCCTCTCTAT 1413
158 GTTACACTGAAGAATGGCGAGTGTCTATCTTTGGTCACTGCCCTCTCTAT 217
158 GTTACACTGAAGAATGGCGAGTGTCTATCTTTGGTCACTGCCCTCTCTAT 217
1414 GGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGAGTATATTACAC 1473
1414 GGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGAGTATATTACAC 1473
218 GGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGAGTATATTACAC 277
218 GGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGAGTATATTACAC 277
1474 ACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTCAGCAGCATTAGGACC 1533
1474 ACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTCAGCAGCATTAGGACC 1533
278 ACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTCAGCAGCATTAGGACC 337
278 ACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTCAGCAGCATTAGGACC 337
1534 AGGGCCCTTATACGTTTCATGTTGCTTACAAAGGCTTTTCAAGTGCCTTCAAGTCCGCTTGA 1593
1534 AGGGCCCTTATACGTTTCATGTTGCTTACAAAGGCTTTTCAAGTGCCTTCAAGTCCGCTTGA 1593
338 AGGGCCCTTATACGTTTCATGTTGCTTACAAAGGCTTTTCAAGTGCCTTCAAGTCCGCTTGA 397
338 AGGGCCCTTATACGTTTCATGTTGCTTACAAAGGCTTTTCAAGTGCCTTCAAGTCCGCTTGA 397
1594 GATGATCTCTACCGATATGATGTGATACCCAGATGGACCATCTTAAAGGACAGCGGA 1653
1594 GATGATCTCTACCGATATGATGTGATACCCAGATGGACCATCTTAAAGGACAGCGGA 1653
398 GATGATCTCTACCGATATGATGTGATACCCAGATGGACCATCTTAAAGGACAGCGGA 457
398 GATGATCTCTACCGATATGATGTGATACCCAGATGGACCATCTTAAAGGACAGCGGA 457
1654 TTTTTCCTTACTTTCACACACAGCTGTGATAGTGAG 1688
1654 TTTTTCCTTACTTTCACACACAGCTGTGATAGTGAG 1688
458 TTTTTCCTTACTTTCACACACAGCTGTGATAGTGAG 492
458 TTTTTCCTTACTTTCACACACAGCTGTGATAGTGAG 492

RESULT 37
BI089332
LOCUS 602853292F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4994735 5',
DEFINITION mRNA sequence.
ACCESSION BI089332
VERSION BI089332.1 GI:14507662
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1017 row: g column: 24
High quality sequence stop: 676.
Location/Qualifiers
1. 706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4994735"

```
/cell_line="MGC36"
/lab_host="DH10B"
/clone_lib="NIH_MGC_10"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match      7.8%; Score 333; DB 12; Length 706;
Best Local Similarity 99.5%; Pred. No. 1.6e-147;
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1537 GCCTATACGTTTCATGTGGCTACAGGCTTTCAGTGCCCAATAGTACCGGCTTGCGAT 1596
Db 47 GCCTATACGTTTCATGTGGCTACAGGCTTTCAGTGCCCAATAGTACCGGCTTGCGAT 106

Qy 1597 GATCTTACCGATATGATGTGGATACCCAGATGTGGACCACTTCTTAAGGACAGCCGATTT 1656
Db 107 GATCTTACCGATATGATGTGGATACCCAGATGTGGACCACTTCTTAAGGACAGCCGATTT 166

Qy 1657 TTCGGTTACTTGCACACAGCTGTGATAGTGTAGTGAGGACCACTGTGGTGTGGGGAAC 1716
Db 167 TTCGGTTACTTGCACACAGCTGTGATAGTGTAGTGAGGACCACTGTGGTGTGGGGAAC 226

Qy 1717 ACACAAATGACACATCTATGAGCCATGCGGCCCAATGCTTCTTTCAGATTTTCATGGCC 1776
Db 227 ACACAAATGACACATCTATGAGCCATGCGGCCCAATGCTTCTTTCAGATTTTCATGGCC 286

Qy 1777 TATGACATTCGCTGTGACCGCTGTGAGTGTCCGACCACTGATCTCCACCACTGATGTC 1836
Db 287 TATGACATTCGCTGTGACCGCTGTGAGTGTCCGACCACTGATCTCCACCACTGATGTC 346

Qy 1837 AACAGATTGGCCATTTCAGACAGTCTTACACAAAGACCACTGATGTGTGGTGGTTTC 1896
Db 347 AACAGATTGGCCATTTCAGACAGTCTTACACAAAGACCACTGATGTGTGGTGGTTTC 406

Qy 1897 AATAGTCTCTCTCAGGACATCCTGTGATATACCTCGGAAACAGTGTGATGCGATCGG 1956
Db 407 AATAGTCTCTCTCAGGACATCCTGTGATATACCTCGGAAACAGTGTGATGCGATCGG 466

Qy 1957 AGTGAAGCCGCTTGT 1971
Db 467 AGTGAAGCCGCTTGT 481

RESULT 38
BI004303 484 bp mRNA linear EST 13-JUN-2001
LOCUS PMO-HN0078-280201-010-h11 HN0078 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI004303
VERSION BI004303.1 GI:14408377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 484)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
PUBMED 10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
```

```
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=PMO&t2=PMO-HN0078-
280201-010-h11&t3=2001-02-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 483.
Location/Qualifiers
1..484
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HN0078"
/notes="Organ: head_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match      7.6%; Score 326; DB 12; Length 484;
Best Local Similarity 99.7%; Pred. No. 3.4e-144;
Matches 376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3545 TTGTGGCTACTCTCAGCAACAAACAGGATTTGGACATGTTTCATCAATGCTTCAAGA 3604
Db 73 TTGTGGCTACTCTCAGCAACAAACAGGATTTGGACATGTTTCATCAATGCTTCAAGA 132

Qy 3605 ATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGAAG 3664
Db 133 ATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGAAG 192

Qy 3665 AGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTA CAAGATAGTTTCTTAATGAGA 3724
Db 193 AGATGCTGCTGTTTCAAAAACCAACATTAAAGAGTA CAAGATAGTTTCTTAATGAGA 252

Qy 3725 AGTTTGATTTTCGCAACCAACCAATATACATTTCTTGTGTTTATGTCAGTAATTTCACT 3784
Db 253 AGTTTGATTTTCGCAACCAACCAATATACATTTCTTGTGTTTATGTCAGTAATTTCACT 312

Qy 3785 GGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAATTTTATGGACCTGGTACAGT 3844
Db 313 GGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAATTTTATGGACCTGGTACAGT 372

Qy 3845 TCTTGAGTACTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3904
Db 373 TCTTGAGTACTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 432

Qy 3905 AGATCAACAAAGTTGT 3921
Db 433 AGATCAACAAAGTTGT 449

RESULT 39
BI004303 399 bp mRNA linear EST 17-JAN-2001
LOCUS QV1-ET0181-021200-519-d10 ET0181 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI004303
VERSION BI004303.1 GI:12271081
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 399)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
```

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matekuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QVt2=QV1-ET0181-021200-519-d10&t3=2000-12-02&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 399.

Location/Qualifiers

1. .399

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="ET0181"

/note="Organ: lung tumor; Vector: puc18; Site: 1: Smal; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES

source

ORIGIN

Query Match 7.2%; Score 309; DB 10; Length 399;
Best Local Similarity 100.0%; Pred. No. 4.5e-136;
Matches 309; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3159 CAATGCTCTTCAATTCACGTGTCAGCTTGCATGCAAGCGGCACAGTAATGCAATCAA 3218

Db 26 CAATGCTCTTCAATTCACGTGTCAGCTTGCATGCAAGCGGCACAGTAATGCAATCAA 85

QY 3219 TCAGAGCATCTGTCAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACCTGGCAT 3278

Db 86 TCAGAGCATCTGTCAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACCTGGCAT 145

QY 3279 ATCTGGCTTTACCGTGATCCCAATGAGGGAATGTGACCAATGCAAGTGCATGG 3338

Db 146 ATCTGGCTTTACCGTGATCCCAATGAGGGAATGTGACCAATGCAAGTGCATGG 205

QY 3339 GCAGCGCTCTGTGCAACCAACACCGGCAAGTGTCTGACACCAAGGCGCTCAA 3398

Db 206 GCAGCGCTCTGTGCAACCAACACCGGCAAGTGTCTGACACCAAGGCGCTCAA 265

QY 3399 GGGGAGCAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGG 3458

Db 266 GGGGAGCAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGG 325

QY 3459 AACATGTTA 3467

Db 326 AACATGTTA 334

RESULT 40

AI818764/c

LOCUS

AI818764 671 bp mRNA linear EST 07-MAR-2000

DEFINITION

Wllie04.x1 NCI CGAP Utl Homo sapiens cDNA clone IMAGE:2424606 3' similar to TR:060295 O60295 KIAA0548 PROTEIN ; contains L1.t3 L1 repetitive element ;, mRNA sequence.

ACCESSION

AI818764.1 GI:5437843

VERSION

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1. (bases 1 to 671)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2162 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 411.

Location/Qualifiers

1. .671

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2424606"

/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"

/lab_host="DH10B"

/clone_lib="NCI-CGAP_Utl"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Life Technologies catalog #:

11538-014"

ORIGIN

Query Match 7.0%; Score 300; DB 9; Length 671;
Best Local Similarity 100.0%; Pred. No. 1e-131;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3502 AGTCATATCCAGGAGATGATCGCTATTACACAGTATCAATTTTGGCTACTCTGAC 3561

Db 517 AGTCATATCCAGGAGATGATCGCTATTACACAGTATCAATTTTGGCTACTCTGAC 458

QY 3562 GACAAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCACTCAACATC 3621

Db 457 GACAAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCACTCAACATC 398

QY 3622 ACCTGGGCTGCCAGTTTCTCAGCTGGACCCAGGCTGGAGAGAGATGCTGTGTTTCA 3681

Db 397 ACCTGGGCTGCCAGTTTCTCAGCTGGACCCAGGCTGGAGAGAGATGCTGTGTTTCA 338

QY 3682 AAAACCAACATTAAGGAGTACAAGATAGTTTCTTAATGAGAGTTTGATTTTGGCAAC 3741

Db 337 AAAACCAACATTAAGGAGTACAAGATAGTTTCTTAATGAGAGTTTGATTTTGGCAAC 278

QY 3742 CACCCAAATATCATTCTTCTTTTATGTCACTAGTAATTTCACTGCGCCATCAAAATTCAG 3801

Db 277 CACCCAAATATCATTCTTCTTTTATGTCACTAGTAATTTCACTGCGCCATCAAAATTCAG 218

RESULT 41

AI372809/c

LOCUS

AI372809 383 bp mRNA linear EST 12-JAN-1999

DEFINITION ES175353 Infant brain, Bento Soares Homo sapiens cDNA clone

```

FUH1B1, mRNA sequence.
ACCESSION AI372809
VERSION AI372809.1 GI:4152675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 383)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geohagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudke,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 93 million nucleotides of cDNA sequence
NATURE 377 (6547 Suppl), 3-174 (1995)
96026280
7566098
Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
EST175339 EST17534
Contact: HGI (Human Gene Index)
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-0200
Fax: (301)-838-0208
Email: hgi@tigr.org.
Location/Qualifiers
1. .383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FUH1B1"
/lab_hosts="E. coli DH5-alpha"
/clone_lib="Infant brain, Bento Soares"
/notes="Vector: BA, M13-derived; Site 1: HindIII; Site 2:
NotI; The infant brain library, constructed by Bento
Soares, Columbia University, was oligo-(dT) primed and
directionally cloned into an M13-derived plasmid using
total brain mRNA from a 72-day old human female afflicted
with spinal muscular atrophy."
ORIGIN
Query Match 7.0%; Score 299; DB 9; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.7e-131;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2462 CCCTTTTGGCTTCCTTACACCCAGAGAGGTAGATTTGCTTAAAGCAGTCGGAA 2521
DB 299 CCCTTTTGGCTTCCTTACACCCAGAGAGGTAGATTTGCTTAAAGCAGTCGGAA 240
QY 2522 TAATGCAGTCATCTCAGAGCATGTCCCAAGCTCACCTTTAACCCCATGGGTGGCTTCGGA 2581
DB 239 TAATGCAGTCATCTCAGAGCATGTCCCAAGCTCACCTTTAACCCCATGGGTGGCTTCGGA 180
QY 2582 AGATCAATGTGCTCTACTGTGGGAAGATATGTCCCATTTACAAATAGTTTACTAC 2641
DB 179 AGATCAATGTGCTCTACTGTGGGAAGATATGTCCCATTTACAAATAGTTTACTAC 120
QY 2642 AGTGGATGCCGTCTGAGCCCAAGTCATGCTGGATCTCTGTGGAAATTTTATCAGAACCCAGTA 2701

```

```

DB 119 ACTGGATCCGCTCTGAGCCCTGAGTGTGGATTCTGGAAATTTTATCAGAACCCAGTA 60
QY 2702 CTCGGGGACTGAAGGCTGCACCTGCATCAACCCCACTCAATGCTAGTCTCTGTGAAGG 2760
DB 59 CTCGGGGACTGAAGGCTGCACCTGCATCAACCCCACTCAATGCTAGTCTCTGTGAAGG 1
RESULT 42
AWS79586 427 bp mRNA linear EST 16-MAR-2000
MRO-HT0209-100100-103-e09 HT0209 Homo sapiens cDNA, mRNA sequence.
DEFINITION AWS79586
ACCESSION AWS79586
VERSION AWS79586.1 GI:7254635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 427)
HCGP http://www.ludwig.org.br/ORESTES
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MRO&t2=NR0-HT0209-
100100-103-e09&t3=2000-01-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 427.
Location/Qualifiers
1. .427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0209"
/notes="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 6.9%; Score 298; DB 10; Length 427;
Best Local Similarity 99.7%; Pred. No. 8.4e-131;
Matches 348; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1299 GTTGACCCCTTAAGGCAAGGAGCAGTATGCAGTGTGTGGCACTCTGCACATTTTAC 1358
DB 63 GTTGACCCCTTAAGGCAAGGAGCAGTATGCAGTGTGTGGCACTCTGCACATTTTAC 122
QY 1359 ACTGAGATGCGCCAGTGGTTCATCTGTCATCTTGTGTCATCTGTCCTCTCTATGGATA 1418
DB 123 ACTGAGATGCGCCAGTGGTTCATCTGTCATCTTGTGTCATCTGTCCTCTCTATGGATA 182
QY 1419 TATAAGCAATGTGCAGGAATATGATTTGGATTAAGAACACATGAGTATATTACACCCA 1478
DB 183 TATAAGCAATGTGCAGGAATATGATTTGGATTAAGAACACATGAGTATATTACACCCA 242
QY 1479 GGGTCCCTTGTGGAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGACAGGCGC 1538
DB 243 GGGTCCCTTGTGGAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGACAGGCGG 302

```

QY 1539 CCTATACGTTTCATGTCGCTTCAAGGCTTTTCAGTGCCTCAATAGTACCGGCTTGCAGATGA 1598
 DB 303 CCTATACGTTTCATGTCGCTTCAAGGCTTTTCAGTGCCTCAATAGTACCGGCTTGCAGATGA 362
 QY 1599 TCTCTACCGATGATGATGGATACCCAGATGTCGACCATCTTAAAGAC 1647
 DB 363 TCTCTACCGATGATGATGGATACCCAGATGTCGACCATCTTAAAGAC 411

RESULT 43
 LOCUS AI372810/c 320 bp mRNA linear EST 12-JAN-1999
 DEFINITION EST175354 Infant brain, Bento Soares Homo sapiens cDNA clone
 FUHIBU1, mRNA sequence.
 ACCESSION AI372810
 VERSION AI372810.1 GI:4152676
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 320)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geonhagen,N.S.,
 Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M.,
 Moreno-Palancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.P., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,S.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 96026280
 7566098
 Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
 EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
 EST175339 EST17534
 Contact: HGI (Human Gene Index)
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-0200
 Fax: (301)-838-0208
 Email: hgi@tigr.org

FEATURES
 Location/Qualifiers
 1..320
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="FUHIBU1"
 /lab_host="E. coli DH5-alpha"
 /clone_lib="Infant brain, Bento Soares"
 Note: The infant brain library, constructed by Bento
 Soares, Columbia University, was oligo-(dT) primed and
 directionally cloned into an M13-derived plasmid using
 total brain mRNA from a 72-day old human female afflicted
 with spinal muscular atrophy."

ORIGIN
 Query Match 6.6%; Score 283; DB 9; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.2e-123;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 CCCTTTTGGCTTCTCTTCAACCCAGAGAAGGTAGAAATTTGTCTTAAAGCAGCTCGAA 2521
 DB 283 CCCTTTTGGCTTCTCTTCAACCCAGAGAAGGTAGAAATTTGTCTTAAAGCAGCTCGAA 224
 QY 2522 TAATCGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCGTGGTCCGCTTGGGA 2581
 DB 223 TAATCGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCGTGGTCCGCTTGGGA 164
 QY 2582 AGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCATTTTACAATAGTTTACTAC 2641
 DB 163 AGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCATTTTACAATAGTTTACTAC 104
 QY 2642 AGTGAATGCCCTCTGAGCCAGTGTGCTGGAATTTTATCAGAACCCAGTA 2701
 DB 103 AGTGAATGCCCTCTGAGCCAGTGTGCTGGAATTTTATCAGAACCCAGTA 44
 QY 2702 CTCGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGG 2744
 DB 43 CTCGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGG 1

RESULT 44
 AA350293/c
 LOCUS EST57512 Infant brain Homo sapiens cDNA 5' end, mRNA sequence.
 DEFINITION AA350293
 ACCESSION AA350293
 VERSION AA350293.1 GI:2002620
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 400)
 Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
 3,400 expressed sequence tags identify diversity of transcripts
 from human brain
 Nat. Genet. 4, 256-267 (1993)
 93364420
 8358434
 Other ESTs: EST57511, THC168226
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavage@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 Location/Qualifiers
 1..400
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):151062"
 /db_xref="taxon:9606"
 /sex="female"
 /dev_stage="infant"
 /clone_lib="Infant brain"
 /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
 Site_2: NotI"

ORIGIN
 Query Match 6.6%; Score 283; DB 9; Length 400;
 Best Local Similarity 99.7%; Pred. No. 1.3e-123;
 Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2761 CTTGCAACACACAGTCTTAACAGTGGCGGACACATGTGCTTGAAGACAGCATGTGA 2820
 DB 334 CTTGCAACACACAGTCTTAACAGTGGCGGACACATGTGCTTGAAGACAGCATGTGA 275
 QY 2821 GATTGCACCGCGGAGCTCTTGAGTGCATGTGTGTCAGACACATCAAGCAGTGTGTGAC 2880


```

Db      274  GATTGCACGCGNAGCTCTGAGTGCACTGGTGCAGCAACATGAAGCAGTGTGTGGAC 215
Qy      2881  TCCAATGCCATATGGGCTCCTCCCTTTGGCCAGTGTATGAATGGTATACATGAGC 2940
Db      214  TCCAATGCCATATGGGCTCCTCCCTTTGGCCAGTGTATGAATGGTATACATGAGC 155
Qy      2941  ACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 3000
Db      154  ACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 95
Qy      3001  GGCTGTGGCTGGTGTACTGATCCAGCAATCTGGCAAGGAAATGTCATAGAGGTTCC 3060
Db      94  GGCTGTGGCTGGTGTACTGATCCAGCAATCTGGCAAGGAAATGTCATAGAGGTTCC 35
Qy      3061  TATAAGGACGAGTGAAGTGCCTTGCAGGCC 3094
Db      34  TATAAGGACGAGTGAAGTGCCTTGCAGGCC 1

RESULT 45
BF853021/c
LOCUS      BF853021      483 bp      mRNA      linear      EST 16-JAN-2001
DEFINITION MR2-EN0093-191200-001-905 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF853021
VERSION     BF853021.1  GI:12240869
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1. (bases 1 to 483)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          MEDLINE
          PUBLISHED 191200-001-G05&t3=2000-12-19&t4=1
          COMMENT Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
          191200-001-G05&t3=2000-12-19&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 483.
          Location/Qualifiers
            1..483
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /dev_stage="Adult"
              /clone_lib="EN0093"
              /note="Organ: lung normal; Vector: puc18; Site_1: Smal;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORSTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the pUC 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
FEATURES
source
ORIGIN

```

```

Query Match      6.6%; Score 281; DB 10; Length 483;
Best Local Similarity 99.5%; Pred. No. 1.2e-122;
Matches 381; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3419  GTGAGGTAGAAAATCGATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATCTCTC 3478
Db      448  GTGAGGTAGAAAATCGATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATCTCTC 389
Qy      3479  TTATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTA 3538
Db      388  TTATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTA 329
Qy      3539  TCAATTTTGTGGTACTCTCTGACGAACAAAACAGGATTTGGACATGTTCAATGCGCT 3598
Db      328  TCAATTTTGTGGTACTCTCTGACGAACAAAACAGGATTTGGACATGTTCAATGCGCT 269
Qy      3599  CCAAGAATTTCAACCTCAACATCACCTGGGTGCCAGTTTCTCAGCTGGAACCCAGGCTG 3658
Db      268  CCAAGAATTTCAACCTCAACATCACCTGGGTGCCAGTTTCTCAGCTGGAACCCAGGCTG 209
Qy      3659  GAGAAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTCTCTA 3718
Db      208  GAGAAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTCTCTA 149
Qy      3719  ATGAGAAGTTTGATTTTGGCAACCAACCAATATCAGTTCTTTTATGTCAGTAATT 3778
Db      148  ATGAGAAGTTTGATTTTGGCAACCAACCAATATCAGTTCTTTTATGTCAGTAATT 89
Qy      3779  TCACCTGGCCCATCAAAATTCAG 3801
Db      88  TCACCTGGCCCATCAAAATTCAG 66

RESULT 46
BG015062/c
LOCUS      BG015062      281 bp      mRNA      linear      EST 24-JAN-2001
DEFINITION QV2-GN0208-221200-579-a02 GN0208 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG015062
VERSION     BG015062.1  GI:12466889
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1. (bases 1 to 281)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          MEDLINE
          PUBLISHED 191200-001-G05&t3=2000-12-22&t4=1
          COMMENT Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-GN0208-
          221200-579-a02&t3=2000-12-22&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 281.
          Location/Qualifiers
            1..281
              /organism="Homo sapiens"
FEATURES
source

```

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/dev_stage="Adult"  
/clone_lib="GNO208"  
/note="Organ: placenta normal; Vector: puc18; Site 1:  
SmaI; Site 2: SmaI; A mini-library was made by cloning  
products derived from ORESTES PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."
```

ORIGIN

```
Query Match 6.3%; Score 270; DB 10; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2e-117;  
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1760 CTTGAGATTCATGGCCCTATGACATTCGCTGTGACCGCTGTGAGTCTGCTTCCCGACCTG 1819  
Db 281 CTTGAGATTCATGGCCCTATGACATTCGCTGTGACCGCTGTGAGTCTGCTTCCCGACCTG 222  
QY 1820 ATCTCCACCATGATGTCAACAGATTGGCCATTGACGAGTCTTACACAAAGCAGCACCATGT 1879  
Db 221 ATCTCCACCATGATGTCAACAGATTGGCCATTGACGAGTCTTACACAAAGCAGCACCATGT 162  
QY 1880 ATGTGTCGGTGGTTCAATAGTCTCTCTCCAGCAGATCTCGTATTCACCTCGGAAC 1939  
Db 161 ATGTGTCGGTGGTTCAATAGTCTCTCTCCAGCAGATCTCGTATTCACCTCGGAAC 102  
QY 1940 AGTGTGATGCCATCGGAGTGAACCGCTGTTTACGAGGAGGACCTGTTATTCGGGTG 1999  
Db 101 AGTGTGATGCCATCGGAGTGAACCGCTGTTTACGAGGAGGACCTGTTATTCGGGTG 42  
QY 2000 TGTGGAACACAGGGTGCCTCTCAGTGTATCT 2029  
Db 41 TGTGGAACACAGGGTGCCTCTCAGTGTATCT 12
```

```
RESULT 47  
AA350292  
LOCUS AA350292 396 bp mRNA linear EST 21-APR-1997  
DEFINITION EST57511 Infant brain Homo sapiens cDNA 3' end, mRNA sequence.  
ACCESSION AA350292  
VERSION AA350292.1 GI:2002619  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Adams, M.D., Kerlavage, A.R., Fields, C. and Venter, J.C.  
3,400 expressed sequence tags identify diversity of transcripts  
from human brain  
Nat. Genet. 4, 256-267 (1993)  
93364420  
8358434  
Other ESTs: EST57512 THC168885  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@igir.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (http://www.tigr.org/tdb/hgi/hgi.html)  
Seq primer: M13-21.  
Location/Qualifiers  
1. .396  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="ATCC (inhost):151062"
```

FEATURES

```
source  
1. .396  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="ATCC (inhost):151062"
```

```
/db_xref="taxon:9606"  
/sex="female"  
/dev_stage="infant"  
/clone_lib="Infant brain"  
/note="Organ: brain; Vector: lafmid BA; Site 1: HindIII;  
Site 2: NotI"
```

ORIGIN

```
Query Match 6.3%; Score 270; DB 9; Length 396;  
Best Local Similarity 99.5%; Pred. No. 2.1e-117;  
Matches 370; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1438 TATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTGCCTTTGTGCAAGG 1497  
Db 1 TATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTGCCTTTGTGCAAGG 60  
QY 1498 GGTTCAGGCCATAGCAGTGTTTACGACCATAGGACCGCCCTATACGTTTCATGGTGGC 1557  
Db 61 GGTTCAGGCCATAGCAGTGTTTACGACCATAGGACCGCCCTATACGTTTCATGGTGGC 120  
QY 1558 TACAAGGGCTTTTCAGTCCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTG 1617  
Db 121 TACAAGGGCTTTTCAGTCCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTG 180  
QY 1618 GATACCCAGATGTGACCAATCTTAAAGACAGCCGATTTTCCGTTACTTGCACACAGCT 1677  
Db 181 GATACCCAGATGTGACCAATCTTAAAGACAGCCGATTTTCCGTTACTTGCACACAGCT 240  
QY 1678 GTGATAGTGAAGTGAACCAATGCTGCTGTTGGGGGAAACACACAAATGACATCTATG 1737  
Db 241 GTGATAGTGAAGTGAACCAATGCTGCTGTTGGGGGAAACACACAAATGACATCTATG 300  
QY 1738 ACCCATGGCGCAAAATGCTTCTTCAGATTCATGGCCTATGACATTCCTGTGACCGC 1797  
Db 301 ACCCATGGCGCAAAATGCTTCTTCAGATTCATGGCCTATGACATTCCTGTGACCGC 360  
QY 1798 TGGTCAGTGCCT 1809  
Db 361 TGGTCAGTGCCT 372
```

RESULT 48

```
BE709733 524 bp mRNA linear EST 12-SEP-2000  
LOCUS IL3-HT0618-150600-177-All HT0618 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BE709733  
ACCESSION BE709733.1 GI:10097998  
VERSION BE709733.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
MEDLINE  
10737800  
PUBMED  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome
```

Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=IL3-HT0618-150
 600-177-All&t3=2000-05-15&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 22
 High quality sequence stop: 447.
 Location/Qualifiers

FEATURES

source
 1. 524
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0618"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 6.3%; Score 269; DB 10; Length 524;
 Best Local Similarity 99.7%; Pred. No. 6.8e-117; Indels 0; Gaps 0;
 Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 481 CTCATTGAAGGACAGCCAAATAGAAATAGAGATTCGTTTCAATCATTTTGTCTACAGAG 540
 Db 524 CTCATTGAAGGACAGCCAAATAGAAATAGAGATTCGTTTCAATCATTTTGTCTACAGAG 465
 QY 541 TGTAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGCACCGCTAGTTGCT 600
 Db 464 TGTAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGCACCGCTAGTTGCT 405
 QY 601 GCATTTAGTGGCTTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTT 660
 Db 404 GCATTTAGTGGCTTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTT 345
 QY 661 GCCATCAGGTTATGCTGCTGCATTTTTTTAGTATGCTGCTATATATTTGACTGGA 720
 Db 344 GCCATCAGGTTATGCTGCTGCATTTTTTTAGTATGCTGCTATATATTTGACTGGA 285
 QY 721 TTTAATATTACTACAGTTTGTATGTTGTCCTCAATTAATGCTCAGGCCGAGGAGAGTGT 780
 Db 284 TTTAATATTACTACAGTTTGTATGTTGTCCTCAATTAATGCTCAGGCCGAGGAGAGTGT 225
 QY 781 AAGATCAGTAATACAGCGA 800
 Db 224 AAGATCAGTAATACAGCGA 205

RESULT 49

BU531905 778 bp mRNA linear EST 13-SEP-2002
 LOCUS
 DEFINITION AGENCOURT 10205588 NIH MGC 126 Homo sapiens cDNA clone
 IMAGE:6556819 5', mRNA sequence.

ACCESSION BU531905 1 GI:22842346

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

COMMENT NIH-MGC http://mgi.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT Contact: Robert Strausberg, Ph.D.

EMAIL Email: cgapbs@mail.nih.gov

TISSUE Tissue Procurement: NCI

CDNA Library Preparation: Michael Brownstein Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LNCM2723 row: n column: 19

High quality sequence stop: 480.

FEATURES

source
 1. 778
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6556819"
 /tissue_type="mixed (pool of 40 RNAs)"
 /lab_host="DH10B (T1-phage-resistant)"

/clone_lib="NIH_MGC_126"
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
 Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was
 prepared from a pool of 40 cell line polyA+ RNAs (bladder
 - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
 kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
 5.2%, ovary - 4%, pharynx - 2.3%, prostate - 4.3%,
 salivary gland - 1.3%, and skin - 2.3%). 5' and 3'

adaptors were used in cloning as follows:

5'-AAGCAGTGTATCAACGAGAGTGGCATTACGGCGG-3' and
 5'-ATTCTAGAGCGCGGCGGACATG-DT(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the 0.5-1
 kb size fraction (other fractions present in NIH_MGC_127
 and NIH_MGC_128). Library created in the laboratory of T.
 Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
 Library."

ORIGIN

Query Match 5.9%; Score 253; DB 13; Length 778;
 Best Local Similarity 100.0%; Pred. No. 3.4e-109; Indels 0; Gaps 0;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3549 GGTCTACTCTGACGACAAACACAGGATTTGGACATGTTTCATCATGCTCCCAAGATTT 3608
 Db 3 GGTCTACTCTGACGACAAACACAGGATTTGGACATGTTTCATCATGCTCCCAAGATTT 62
 QY 3609 CAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGAT 3668
 Db 63 CAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGAT 122
 QY 3669 GCCTGTGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTCTCTTAATGAGAAGTT 3728
 Db 123 GCCTGTGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTCTCTTAATGAGAAGTT 182
 QY 3729 TGATTTTCGCAACCCACCAATATCATCTTTCTTTTATGTCAGTAATTTCCCTGGCC 3788
 Db 183 TGATTTTCGCAACCCACCAATATCATCTTTCTTTTATGTCAGTAATTTCCCTGGCC 242
 QY 3789 CATCAAAATTCAG 3801
 Db 243 CATCAAAATTCAG 255

RESULT 50

AI818750/c

LOCUS

DEFINITION

AI818750

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AI818750 589 bp mRNA linear EST 07-MAR-2000
 willc03.x1 NCI CGAP Utl1 Homo sapiens cDNA clone IMAGE:2424580 3'
 similar to TR:O60295 O60295 KIAA0548 PROTEIN ; contains L1.t3 L1
 repetitive element ;, mRNA sequence.

AI818750.1 GI:5437829

EST.

Homo sapiens

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 589)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Cloned distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2056 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 399.
Location/Qualifiers
source
1. 589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2424580"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_hosts="DH10B"
/clone_lib="NCI CGAP Ut1"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

ORIGIN
Query Match 5.8%; Score 249; DB 9; Length 589;
Best Local Similarity 99.7%; Pred. No. 2.6e-107;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3502 AGTCTATCCGAGGAGATGCGTATATACAGCTATCAATTTGGCTACTCTCGAC 3561
DB |-----|
QY 517 AGTCTATCCGAGGAGATGCGTATATACAGCTATCAATTTGGCTACTCTCGAC 458
DB |-----|
QY 3562 GAAACAACACAGGATTTGGACATGTTTCATCAATCCCTCCAGAAATTTCAACCTCAACATC 3621
DB |-----|
QY 457 GAAACAACACAGGATTTGGACATGTTTCATCAATCCCTCCAGAAATTTCAACCTCAACATC 398
DB |-----|
QY 3622 ACCTGGGCTGCCATTTCTCAGCTGGAACCCAGGCTCGAAGAGATGCGCTGTGTTTCA 3681
DB |-----|
QY 397 ACCTGGGCTGCCATTTCTCAGCTGGAACCCAGGCTCGAAGAGATGCGCTGTGTTTCA 338
DB |-----|
QY 3682 AAAACCAACATTAAGGAGTCAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAAC 3741
DB |-----|
QY 337 AAAACCAACATTAAGGAGTCAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAAC 278
DB |-----|
QY 3742 CACCCAAATACACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAG 3801
DB |-----|
QY 277 CACCCAAATACACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAG 218
DB |-----|

RESULT 51
BF928756
LOCUS BF928756 313 bp mRNA linear EST 19-JAN-2001
DEFINITION IL5-NT0228-061200-308-f12 NT0228 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF928756
VERSION BF928756.1 GI:12326884
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 313)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., D.H.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,

O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL5&t=IL5-NT0228-
061200-308-f12&t3=2000-12-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 313.
Location/Qualifiers
source
1. 313
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0228"
/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 5.7%; Score 243; DB 10; Length 313;
Best Local Similarity 99.7%; Pred. No. 1.7e-104;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3307 GGAGGGAATGTCAGCCATGCAAGTGCATGGGACGCGTCTCTGCAACACCAACACG 3366
DB |-----|
QY 7 GGAGGGAATGTCAGCCATGCAAGTGCATGGGACGCGTCTCTGCAACACCAACACG 66
DB |-----|
QY 3367 GGCAAGTCTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTA 3426
DB |-----|
QY 67 GGCAAGTCTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTA 126
DB |-----|
QY 3427 GAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTTATTGAC 3486
DB |-----|
QY 127 GAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTTATTGAC 186
DB |-----|
QY 3487 TATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTT 3546
DB |-----|
QY 197 TATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTT 246
DB |-----|
QY 3547 GTGGCTACTCTGACGAACAAACAGGATTGGACATGTTTCATCAATGCTCTCC 3600
DB |-----|
QY 247 GTGGCTACTCTGACGAACAAACAGGATTGGACATGTTTCATCAATGCTCTCC 300
DB |-----|

RESULT 52
R87660
LOCUS R87660 416 bp mRNA linear EST 16-AUG-1995
DEFINITION yp95g01.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:194640 3', mRNA sequence.
ACCESSION R87660
VERSION R87660.1 GI:946473
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoideae; Homo.
1 (bases 1 to 416)
REFERENCE
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,P.
TITLE
The WashU-Merck EST Project
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1984
High quality sequence stops: 273
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1984 Std Error: 0.00
Seq primer: Promega -2lm3
High quality sequence stop: 273.
Location/Qualifiers
1. 416
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="GDB:3763690"
/db_xref="taxon:9606"
/clone="IMAGE:194640"
/sex="male"
/dev stage="20 week-post conception fetus"
/lab host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: pVT73D (Pharmacia)
with a modified polylinker, Site_1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' - AACTGGAGAATAATTAAGACATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pVT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaudo."

ORIGIN

Query Match	5.6%	Score 241;	DB 14;	Length 416;
Best Local Similarity	100.0%;	Pred. No. 1.6e-103;		
Matches 241;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 2257 ATGTACTACTGTACACAGACACGCTGCAGAGCTGTCCCTGGACACGAACTGCCAG 2316
1 ATGTACTACTGTACACAGACACGCTGCAGAGCTGTCCCTGGACACGAACTGCCAG 60

QY 2317 TGGGAGCCCCGGAACTCAGGAGTGCATTGTCCTGCCGAAATATCTGTGCATTGGCTGG 2376
61 TGGGAGCCCCGGAACTCAGGAGTGCATTGTCCTGCCGAAATATCTGTGCATTGGCTGG 120

QY 2377 CATTTGGTTGGAACCTCATGCTTTGAAATTTACTACTGCCAAGAGGAATTTAGCAATGCT 2436
121 CATTTGGTTGGAACCTCATGCTTTGAAATTTACTACTGCCAAGAGGAATTTAGCAATGCT 180

QY 2437 AAATTGTCTGTAGGAACCAATATGCCCTTTTGGCTTCTCTTCAACCCAGAGAAGGTA 2496
181 AAATTGTCTGTAGGAACCAATATGCCCTTTTGGCTTCTCTTACACCCAGAGAAGGTA 240

QY 2497 G 2497
241 G 241

RESULT 53 .			
BE671400/c			
LOCUS	BE671400	445 bp	linear
		mRNA	EST 08-SEP-2000

DEFINITION	7e50d11.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3285909 3' similar to TR:060295 KIA0548 PROTEIN ; contains 11.b1 L1 repetitive element ;, mRNA sequence.		
ACCESSION	BE671400		
VERSION	BE671400.1		
KEYWORDS	GI:10031941		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 445)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaps-r@mail.nih.gov		
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D.		
	CDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: infoimage.llnl.gov		
	Seq primer: -40UP from Gibco		
	High quality sequence stop: 397.		
FEATURES	Location/Qualifiers		
source	1..445		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3285909"		
	/tissue_type="carcinoid"		
	/lab_host="DH10B"		
	/clone_lib="NCI_CGAP_Lu24"		
	/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."		
ORIGIN			
Query Match	5.2%; Score 225; DB 10; Length 445;		
Best Local Similarity	100.0%; Pred. No. 7.7e-96;		
Matches	225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	3577	TTTGACATGTTTCATCAATGCCTCCAGAATTTTCAACCTCAACATCACTGGGCTGCCAGT	3636
Db	445	TTTGACATGTTTCATCAATGCCTCCAGAATTTTCAACCTCAACATCACTGGGCTGCCAGT	386
Qy	3637	TTTCTAGCTGGAAACCCAGGCTGGAGAAGATGCCTGTTGTTTCAAAAACCAACATTAA	3696
Db	385	TTTCTAGCTGGAAACCCAGGCTGGAGAAGATGCCTGTTGTTTCAAAAACCAACATTAA	326
Qy	3697	GAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACCCAAATATCACT	3756
Db	325	GAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACCCAAATATCACT	266
Qy	3757	TTCTTTGTTTATGTCAGTAAATTTTCACCTGCCCATCAAAATTGAG	3801
Db	265	TTCTTTGTTTATGTCAGTAAATTTTCACCTGCCCATCAAAATTGAG	221

RESULT 54	AI076479/c	LOCUS	AI076479	464 bp	mRNA	linear	EST 01-OCT-1998
DEFINITION	oz28b04.x1	Soares_total_fetus_nb2HF8_9w	Homo sapiens	cDNA clone	IMAGE:1676923	3', similar to contains L1.b1 L1 repetitive element	

mRNA sequence.
 AI076479
 VERSION
 EST.
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 464)
 NCBI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1450 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 444.
 FEATURES
 Location/Qualifiers
 1..464
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1676623"
 /dev_stage="8-9 weeks"
 /lab_host="DH108"
 /clone_lib="Soares total fetus Nb2HF8 9w"
 /note="Vector: pT73-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGCCGCTTAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "
 ORIGIN
 Query Match 5.2%; Score 223; DB 9; Length 464;
 Best Local Similarity 100.0%; Pred. No. 7e-95;
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3579 GGACATGTTTCATCAATGCTCCCAAGATTCAACCTCAACATCACTGGGCTGCCAGTTT 3638
 Db 464 GGACATGTTTCATCAATGCTCCCAAGATTCAACCTCAACATCACTGGGCTGCCAGTTT 405
 QY 3639 CTCAGTGGAAACCCAGGCTGGAGAGAGATGCTTGTCTTTCAAAACCAACATTAAAGGA 3698
 Db 404 CTCAGTGGAAACCCAGGCTGGAGAGAGATGCTTGTCTTTCAAAACCAACATTAAAGGA 345
 QY 3699 GTACAAGATAGTTTCTCTATGAGAGTTGATTTTCGCAACCAACCAATATCACTTT 3758
 Db 344 GTACAAGATAGTTTCTCTATGAGAGTTGATTTTCGCAACCAACCAATATCACTTT 285
 QY 3759 CTTTGTGTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 3801
 Db 284 CTTTGTGTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 242
 RESULT 55
 BX644466
 LOCUS
 DEFINITION
 DKF2p781N2432 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
 EST.
 ACCESION
 BX644466
 VERSION
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 517)
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevas, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
 and Marra, M.

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 500)
 Bloeker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 Pobo, G., Han, M., and Wiemann, S.
 EST (Bloeker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., et al.)
 Unpublished (2003)
 Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GfP (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.
 This clone (DKFZp781N2432) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de.
 FEATURES
 Location/Qualifiers
 1..500
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp781N2432"
 /dev_stage="adult"
 /lab_host="DH108"
 /clone_lib="781 (synonym: hlcc4)"
 /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
 cDNA-collection"
 ORIGIN
 Query Match 4.3%; Score 185; DB 13; Length 500;
 Best Local Similarity 99.6%; Pred. No. 1.1e-76;
 Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4055 TTCCCAAAACCCATTGACATGGAGCCGCTGTTTGGCAACAAAGCCGCTCTCTGTGT 4114
 Db 1 TTCCCAAAACCCATTGACATGGAGCCGCTGTTTGGCAACAAAGCCGCTCTCTGTGT 60
 QY 4115 TTGTGAGGCTCCCTCGAGGCTGGGGCATCCCTCTCTCGGCGAGTCAGTCTGTGTG 4174
 Db 61 GTGTGAGGCTCCCTCGAGGCTGGGGCATCCCTCTCTCGGCGAGTCAGTCTGTGTG 120
 QY 4175 TGCCACGCGCCCTGTGTGGACATTTCTCAGCAGATCCGATAGTGTACAGGAGAGTCAG 4234
 Db 121 TGCCACGCGCCCTGTGTGGACATTTCTCAGCAGATCCGATAGTGTACAGGAGAGTCAG 180
 QY 4235 GAGCCGTGAGAAACCGAGAGCAGCAGCCCTTGCACAGCCTGGACCTGCATCTGA 4290
 Db 181 GAGCCGTGAGAAACCGAGAGCAGCAGCCCTTGCACAGCCTGGACCTGCATCTGA 236
 RESULT 56
 T69367
 LOCUS
 DEFINITION
 yc37g02.r1 Stratagene liver (#937224) Homo sapiens cDNA clone
 IMAGE:82898 5', mRNA sequence.
 ACCESION
 T69367
 VERSION
 T69367.1 GI:680515
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 517)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevas, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
 and Marra, M.


```

|||||
10 AGAATTTGTCCTTAAGCAGCTGGAAATATGCAAGTCAATCTCAGAGCATGTCACAGCTCAC 69
|||||
2556 CTTAACCCCGATGGTCGGCTTCGGAGATCAATGTCCTACTGTCGCTGGGAGATAT 2615
|||||
70 CTTAACCCCGATGGTCGGCTTCGGAGATCAATGTCCTACTGTCGCTGGGAGATAT 129
|||||
2616 GTCCCCATTTACAATATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATT 2675
|||||
130 GTCCCCATTTACAATATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATT 189
|||||
2676 CTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 2711
|||||
190 CTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 225

```

```

RESULT 61
LOCUS T92696 355 bp mRNA linear EST 22-MAR-1995
DEFINITION ye26f05.11 Stratagene lung (#937210) Homo sapiens cDNA clone
IMAGE:118881 5', mRNA sequence.
ACCESSION T92696
VERSION T92696.1 GI:724609
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 355)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierrey-Meg, J.,
Trevasakis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478
8889549
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 271
High quality sequence stops: 255 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 271 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 255.

```

```

FEATURES
source
1..355
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:487170"
/db_xref="taxon:9606"
/clone="IMAGE:118881"
/sex="male"
/dev_stages="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene lung (#937210)"
/notes="Organ: lung; Vector: pBluescript SK-; Site:1:
ECORI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGACGAG
3' -3' adaptor sequence: 5' CTCAGATTTTTTTTTTTTTTTT 3'"

```

```

ORIGIN
Query Match 3.2%; Score 136; DB 14; Length 355;
Best Local Similarity 100.0%; Pred. No. 3e-53;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
3667 ATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAG 3726
|||||

```

```

Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2568 GGTGCGCCTTCGGAAGATCAATGTCCTACTGTCGCTGGGAGATATGTCCTATTTAC 2627
DB 73 GGTGCGCCTTCGGAAGATCAATGTCCTACTGTCGCTGGGAGATATGTCCTATTTAC 132
QY 2628 AAATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATTTT 2687
DB 133 AAATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATTTT 192
QY 2688 ATCAGAACCCAGTACT 2703
DB 193 ATCAGAACCCAGTACT 208

```

```

RESULT 62
LOCUS BF475948/c 413 bp mRNA linear EST 05-DEC-2000
DEFINITION nac42h02.xl Lupski sciatic nerve Homo sapiens cDNA clone
IMAGE:3395955 3', similar to TR:060295 KIAA0548 PROTEIN
; contains LI.D1 LI repetitive element ;, mRNA sequence.
ACCESSION BF475948
VERSION BF475948.1 GI:11546775
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 413)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps@mail.nih.gov
Tissue Procurement: Dr. James Lupski
cDNA Library Preparation: Lupski Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 364.

```

```

FEATURES
source
1..413
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3395955"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stages="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site:1:
NotI; Site:2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCGCG-3' and
5'-GACTACTTCTAGATCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

```

```

ORIGIN
Query Match 3.1%; Score 135; DB 10; Length 413;
Best Local Similarity 100.0%; Pred. No. 9.4e-53;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
3667 ATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAG 3726
|||||

```

Db 377 ATGCGCTGTTGTTTCAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAATGAGAG 318
 QY 3727 TTGTATTTTCGCAACACCAACCAATATCATCTTTGTTTGTATGTCAGTAATTCACCTGG 3786
 Db 317 TTGTATTTTCGCAACACCAACCAATATCATCTTTGTTTGTATGTCAGTAATTCACCTGG 258
 QY 3787 CCCATCAAAATTCAG 3801
 Db 257 CCCATCAAAATTCAG 243

RESULT 63
 AA297214
 LOCUS
 DEFINITION EST112757 Embryo, 12 week I Homo sapiens cDNA 5' end, mRNA

ACCESSION
 VERSION AA297214.1 GI:1949548
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 (bases 1 to 279)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Frichman,J.L., Geoghegan,N.S.,
 Glodek,A., Gnehm,C.B., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Gao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dinke,D., Peng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Hasseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Other ESTs: THCI68226
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tadb/hgi.html>)

Seq primer: M3 Reverse.
 Location/Qualifiers
 1..279
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):129808"
 /db_xref="taxon:9606"
 /dev_stage="embryo, 12 wks"
 /clone_lib="Embryo, 12 week 1"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN
 Query Match 3.0%; Score 129; DB 9; Length 279;
 Best Local Similarity 100.0%; Pred. No. 6.4e-50;
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2979 TAGTCATTGCTTGGAGCAACCAAGCGTGTGGCTGGTGTACTGATCCAGCAATCTGGCAA 3038
 Db 42 TAGTCATTGCTTGGAGCAACCAAGCGTGTGGCTGGTGTACTGATCCAGCAATCTGGCAA 101
 QY 3039 AGGGAATGCATAGAGGGTTCTTATTAAGGACACAGTGAAGATGCTTCGCAAGCCCTTAC 3098
 Db 102 AGGGAATGCATAGAGGGTTCTTATTAAGGACACAGTGAAGATGCTTCGCAAGCCCTTAC 161
 QY 3099 AGGAATTT 3107
 Db 162 AGGAATTT 170

RESULT 64
 AA771958
 LOCUS
 DEFINITION ai37e06.s1 Soares parathyroid tumor, NBHPA Homo sapiens cDNA clone
 1359202 3', similar to contains L1.t3 L1 repetitive element ;, mRNA
 sequence.

ACCESSION
 VERSION AA771958.1 GI:2823741
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 (bases 1 to 370)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 895 Std Error: 0.00
 Seq primer: -40m3 fwd. Et from Amersham
 High quality sequence stop: 297.
 Location/Qualifiers
 1..370
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="1359202"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares parathyroid tumor NBHPA"
 /note="Organ: parathyroid gland; Vector: pT7T3D
 (Pharmacia) with a modified polylinker; Site_1: Not I;
 Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGGCGCGCCGACCAATTTTTTTTTTTTTT
 TTTT-3'] double-stranded cDNA was size selected, ligated
 to Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT7T3
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

FEATURES
 source
 1..370
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="1359202"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares parathyroid tumor NBHPA"
 /note="Organ: parathyroid gland; Vector: pT7T3D
 (Pharmacia) with a modified polylinker; Site_1: Not I;
 Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGGCGCGCCGACCAATTTTTTTTTTTTTT
 TTTT-3'] double-stranded cDNA was size selected, ligated
 to Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT7T3
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

ORIGIN
 Query Match 2.9%; Score 126; DB 9; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.9e-48;
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3676 GTTTCACCAACCAATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTT 3735
Db 370 GTTTCACCAACCAATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTT 311
QY 3736 CGCAACCAACCAATTAATCACTTCTTTGTTTATGTCAGTAATTTCACTGGGCCCAATCAAN 3795
Db 310 CGCAACCAACCAATTAATCACTTCTTTGTTTATGTCAGTAATTTCACTGGGCCCAATCAAN 251
QY 3796 ATTTCAG 3801
Db 250 ATTTCAG 245

RESULT 65
AA298732 319 bp mRNA linear EST 18-APR-1997
LOCUS EST114356 Supt cells Homo sapiens cDNA 5' end similar to similar to
DEFINITION laminin, B1 chain, mRNA sequence.
ACCESSION AA298732
VERSION AA298732.1 GI:1951085
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 319)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
Shall,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Peng,D.-F., Perrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other ESTs: THC168226
Contact: Kerlavage, AR
Bioinformatics The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699423
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 319
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):184690"
/db_xref="taxon:9606"
/cell_type="T-lymphocyte"
/clone_lib="Supt cells"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

FEATURES
source
```

```
Query Match 2.9%; Score 125; DB 9; Length 319;
Best Local Similarity 99.9%; Pred. No. 5.5e-48; Indels 0; Gaps 0;
Matches 275; Conservative 0; Mismatches 3;
QY 3016 ACTGATCCCGAGCAATACTGGCAAGGGAATGTCATAGAGGGTTCCTATATAAGGACCAAGTG 3075
Db 1 ACTGATCCCGAGCAATACTGGCAAGGGAATGTCATAGAGGGTTCCTATATAAGGACCAAGTG 60
QY 3076 AAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTCAGC 3135
Db 61 AAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTCAGC 120
QY 3136 ATGTGTCTAGAGGACAGACATCAACTGGTCTTTTCACTTCAGCTTCAGCTTCAGCAATGC 3195
Db 121 ATGTGTCTAGAGGACAGACATCAACTGGTCTTTTCACTTCAGCTTCAGCTTCAGCAATGC 180
QY 3196 AACGGCCACAGTAATGATCAATCAGAGCACTCTGTGAGAAGTGTGAGACCTGACACCA 3255
Db 181 AACGGCCACAGTAATGATCAATCAGAGCACTCTGTGAGAAGTGTGAGACCTGACACCA 240
QY 3256 GCGAAGCACTCGAGACCTGCATATCTGGCTTCTACGG 3293
Db 241 GCGAAGCACTCGAGACCTGCATATCTGGCTTCTACGG 278

RESULT 66
AA297697 230 bp mRNA linear EST 18-APR-1997
LOCUS EST113254 Fetal heart II Homo sapiens cDNA 5' end, mRNA sequence.
DEFINITION AA297697
ACCESSION AA297697.1 GI:1950031
VERSION AA297697.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 230)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
Shall,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Peng,D.-F., Perrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other ESTs: THC168226
Contact: Kerlavage, AR
Bioinformatics The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699423
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 230
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):184690"
/db_xref="taxon:9606"
/cell_type="T-lymphocyte"
/clone_lib="Supt cells"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

FEATURES
source
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):149184"
/db_xref="taxon:9606"
/dev_stage="fetus"
/clone_lib="Fetal heart II"
/notes="Organ: heart; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Query Match      2.9%; Score 123; DB 9; Length 230;
Best Local Similarity 100.0%; Pred. No. 4.5e-47;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2750 TCTGTGAAGGCTTCACCAACACAGTGTCTAGACGAGTCCGACACCATGTCCTTGGAGGA 2809
Db 14 TCTGTGAAGGCTTCACCAACACAGTGTCTAGACGAGTCCGACACCATGTCCTTGGAGGA 73

QY 2810 CAGCATGTGGAGATTGCACACGAGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGC 2869
Db 74 CAGCATGTGGAGATTGCACACGAGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGC 133

QY 2870 AGT 2872
Db 134 AGT 136

RESULT 67
BF700035 683 bp mRNA linear EST 22-DEC-2000
DEFINITION 602127513F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4284387 5',
LOCUS mRNA sequence.
ACCESSION BF700035
VERSION BF700035.1 GI:11985443
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 683)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM117 row: j column: 04
High quality sequence stop: 615.
FEATURES
Location/Qualifiers
1..683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4284387"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_56"
/notes="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgccg); Site 2: SfiI (ggcattatggcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech

```

```

Laboratories (Palo Alto, CA)."

ORIGIN
Query Match      2.9%; Score 123; DB 10; Length 683;
Best Local Similarity 100.0%; Pred. No. 5.9e-47;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2836 AGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGCAATGCCCTATGTG 2895
Db 202 AGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGCAATGCCCTATGTG 261

QY 2896 GCTCTCTCCCTTTGGCCAGTGTATGNAATGATACATGAGCAGCTGCCCTCTGAA 2955
Db 262 GCTCTCTCCCTTTGGCCAGTGTATGNAATGATACATGAGCAGCTGCCCTCTGAA 321

QY 2956 AAT 2958
Db 322 AAT 324

RESULT 68
BF700035 837 bp DNA linear GSS 08-JUN-2003
DEFINITION WHAAD22TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
LOCUS sapiens genomic clone MCF7_1-1D20, genomic survey sequence.
ACCESSION BF700035
VERSION BF700035.1 GI:31519133
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q.,
Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
Gray, J.W. and Collins, C.
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
MEDLINE 22709111
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES
Location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-1D20"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7.1)"
/notes="vector: pECBAC1; Site_1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."

ORIGIN
Query Match      2.8%; Score 119; DB 28; Length 837;
Best Local Similarity 100.0%; Pred. No. 5e-45;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4049 AGACTGTTCACCAACCATTCGACTGGAGCGGTGTTTGCAACAAAGCCGCTGTCTCT 4108
Db 36 AGACTGTTCACCAACCATTCGACTGGAGCGGTGTTTGCAACAAAGCCGCTGTCTCT 95

```

QY 4109 CTGTGTTTGTGAGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCTGGGCACTCAGGT 4167
 Db 96 CTGTGTTTGTGAGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCTGGGCACTCAGGT 154

RESULT 69
 BG945444/c
 LOCUS BG945444 134 bp mRNA linear EST 11-JUN-2001
 DEFINITION P2-AN0093-190401-008-h10 AN0093 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG945444
 VERSION BG945444.1 GI:14344815
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 134)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F.F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.J.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&t2=P2-AN0093-190401-008-h10&t3=2001-04-19&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 33
 High quality sequence stop: 134.
 Location/Qualifiers
 1..134
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="AN0093"
 /note="Organ: amnion normal; Vector: puc18; Site_1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Query Match 2.7%; Score 115; DB 12; Length 134;
 Best Local Similarity 100.0%; Pred. No. 2.8e-43;
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2606 GGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCGCTGAGCCCAAGTG 2665
 Db 134 GGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCGCTGAGCCCAAGTG 75

QY 2666 ATGCTGGATTCTGTGGATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTGC 2720
 Db 74 ATGCTGGATTCTGTGGATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTGC 20

RESULT 70

BX669805 530 bp mRNA linear EST 24-OCT-2003
 EX669805 Sus Scrofa library (scac) Sus scrofa cDNA clone
 scac00291.b.13 3prim, mRNA sequence.
 BX669805
 VERSION BX669805.1 GI:37979704
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 530)
 AUTHORS Bonnet, A., Tosser-Klopp, G., Benne, F., Cabau, C., Villegier, S., Soares, M., Bonaldi, F. and Hately, F.
 TITLE A Pig Normalised Multi-Tissue cDNA Library
 JOURNAL Unpublished (2003)
 COMMENT Contact: Tosser-Klopp G
 Genetique Animale
 Institut National de la Recherche Agronomique
 Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan cedex, FRANCE
 Tel: 33 (0) 5.61.28.51.14
 Fax: 33 (0) 5.61.28.53.08
 Email: tosser@toulouse.inra.fr
 Clone distribution: AGENAE Resource centre, Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du Genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73
 Sequence cleaned of vector, adaptor and repetitions. Contact us at sigenasupport@jouy.inra.fr to obtain the chromatogram of this sequence.
 Plate: 0029 row: b column: 13.
 Location/Qualifiers
 1..530
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="scac00291.b.13"
 /tissue_type="mixed"
 /clone_lib="Sus Scrofa library (scac)"
 /note="Vector: p773D-pac vector; tissues: adipose tissue, brain, kidney, liver, muscle, ovary, testis, heart, hypothalamus, pancreas, skin, spleen, thymus, placenta, pituitary gland, seminal vesicle, small intestine, uterus, adrenals, bulbo uretral gland, cerebral trunk, epididymis, female gonad, gall-bladder, hippocampus, large intestine, male gonad, melanocytes, stomach, udder"

ORIGIN
 Query Match 2.6%; Score 113; DB 13; Length 530;
 Best Local Similarity 100.0%; Pred. No. 3.4e-42;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 478 TGCTCATTTGAAGGACGACCAATAGATPAATGAGACTTCGTTTCATCATTTGTACA 537
 Db 387 TGCTCATTTGAAGGACGACCAATAGATPAATGAGACTTCGTTTCATCATTTGTACA 328

QY 538 GAGTGTAGTGGGACCAATATATGTTTATGTTGGGACTCAATTTATGCACC 590
 Db 327 GAGTGTAGTGGGACCAATATATGTTTATGTTGGGACTCAATTTATGCACC 275

RESULT 71
 CE218575/c
 LOCUS CE218575 363 bp DNA linear GSS 25-SBP-2003
 DEFINITION tigr-gss-dog-17000326860416 Dog Library Canis familiaris genomic, genomic survey sequence.
 ACCESSION CE218575
 VERSION CE218575.1 GI:35374244
 KEYWORDS GSS.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 363)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES
source
1.363
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 2.6%; Score 110; DB 29; Length 363;
Best Local Similarity 100.0%; Pred. No. 8.5e-41;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2242 CCCAGGTAACCCCATGCTACTACTGTACAGAGACCCAGCTCCAGAGCTGCGCCGTG 2301
Db 197 CCCAGGTAACCCCATGCTACTACTGTACAGAGACCCAGCTCCAGAGCTGCGCCGTG 138
QY 2302 GACCAGAACTCCAGTGGAGCCCGGAATCAGGAGTGCAATGCGCCGCC 2351
Db 137 GACCAGAACTCCAGTGGAGCCCGGAATCAGGAGTGCAATGCGCCGCC 88

RESULT 72
BF316716
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BF316716 942 bp mRNA linear EST 21-NOV-2000
Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 363)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 363)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES
source
1.363
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 2.5%; Score 106; DB 10; Length 942;
Best Local Similarity 99.4%; Pred. No. 8.7e-39;
Matches 156; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1238 CAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCATGGTGT 1297
Db 57 CAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCATGGTGT 116
QY 1298 TGTTCACCCCTAAGCAAGGAGGAGTATGCAAGTGGTGGGCACTCTGCACACATTTGTA 1357
Db 117 TGTTCACCCCTAAGCAAGGAGGAGTATGCAAGTGGTGGGCACTCTGCACACATTTGTA 176
QY 1358 CAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCATGGTGT 1394
Db 177 CAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCATGGTGT 213

RESULT 73
BQ307981
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BQ307981 496 bp mRNA linear EST 16-MAY-2002
Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 496)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 496)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 496)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

QY 3323 CATGCAAGTGCAATGGGCACCGCTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCA 3382
DB |||||
340 CATGCAAGTGCAATGGGCACCGCTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCA 281
DB |||||
QY 3383 CCACCAAGGGCGTCAAGGGGACGAGTGCCAGCT 3416
DB |||||
280 CCACCAAGGGCGTCAAGGGGACGAGTGCCAGCT 247
DB |||||

Search completed: March 1, 2004, 20:45:14
Job time : 9548 secs

C 97 24 0.6 345 4 ABL21343 Drosophil
C 98 24 0.6 409 5 ABV03540 Human pro
C 99 24 0.6 412 5 ABV33838 Human pro
C 100 24 0.6 412 5 ABV42727 Human pro

ALIGNMENTS

RESULT 1

AAZ52276
ID AAZ52276 standard; cDNA; 4290 BP.

XX

AC

AAZ52276;

XX

DT 18-JUL-2000 (first entry)

XX

DE Human membrane attractin-2 cDNA.

XX

KW Human; membrane attractin-2; immune response; macrophage; monocyte;

XX

KW T cell; immunostimulant; immunosuppressed patient; cancer;

XX

KW immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

PH CDS

1..4290

FT /*tag= a

FT /product= "Membrane attractin-2"

XX

PN WO200015651-A1.

XX

PD 23-MAR-2000.

XX

PF 14-SEP-1999; 99WO-US020948.

XX

PR 14-SEP-1998; 98US-0100137P.

XX

PA (DAND) DANA FARBER CANCER INST INC.

XX

PI Duke-Cohan JS, Schlossman SF;

XX

DR WPI: 2000-271373/23.

XX

DR P-PDB; AAY70691.

XX

PT Isolated nucleic acids encoding human attractin polypeptides useful for

XX

PS enhancing immune responses.

XX

PS Claim 3; Fig 14; 120pp; English.

XX

CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human membrane attractin-2.
CC Membrane attractin differs from soluble attractin in having a
CC transmembrane domain and a cytoplasmic domain at the C-terminal.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease

XX

SQ Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;

XX

Query Match 100.0%; Score 4290; DB 3; Length 4290;

XX

Best Local Similarity 100.0%; Pred. No. 0;

XX

Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGGCCGAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACCGGCGGCGAGCGCA 60
DB 1 ATGGTGGCCGAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACCGGCGGCGAGCGCA 60
QY 61 GGGCTCGGGGAGGAGCGGCGGCGGCACTGGGACTGGGAGCTGAGGAGGAGGAGGAGGAGG 120
DB 61 GGGCTCGGGGAGGAGCGGCGGCGGCACTGGGACTGGGAGCTGAGGAGGAGGAGGAGGAGG 120
QY 121 CCGGGGCTGGGGGCGGGGCTGGGCTCCGGCGGCTGGGCTCCGACCGGCGGCGGCGGCGG 180
DB 121 CCGGGGCTGGGGGCGGGGCTGGGCTCCGGCGGCTGGGCTCCGACCGGCGGCGGCGGCGG 180
QY 181 CTG 240
DB 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 GCGGAGCGCGGCGGCGGCGGCGGCGGCTGGGCTGAGCCGCGGCGGCGGCGGCGGCGG 300
DB 241 GCGGAGCGCGGCGGCGGCGGCGGCGGCTGGGCTGAGCCGCGGCGGCGGCGGCGGCGG 300
QY 301 TGTGACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 TGTGACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CCGCGCGGCTGGGCTGGGCGGAGCAATGCCAGCACTGCGGGGGCGGCTTCCAGCTAACTGGA 420
DB 361 CCGCGCGGCTGGGCTGGGCGGAGCAATGCCAGCACTGCGGGGGCGGCTTCCAGCTAACTGGA 420
QY 421 TCTTCTGGGTTTGTGACAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 TCTTCTGGGTTTGTGACAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 CTCAATTGAAGGACAGCCAAATAGAATAGAGACTTCTGCTTCAATCAATTTTCTGCTACAGAG 540
DB 481 CTCAATTGAAGGACAGCCAAATAGAATAGAGACTTCTGCTTCAATCAATTTTCTGCTACAGAG 540
QY 541 TGTAGTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGGCTAGTTGCT 600
DB 541 TGTAGTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGGCTAGTTGCT 600
QY 601 GCATTTAGTGGGCTCATTTCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660
DB 601 GCATTTAGTGGGCTCATTTCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660
QY 661 GGCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 GGCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 TTTAATATTACTTACAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 TTTAATATTACTTACAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTAATGTTTGAATAAAGTGAAGTGA 840
DB 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTAATGTTTGAATAAAGTGAAGTGA 840
QY 841 GCATGTCATTTCTCCTCAGTACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 GCATGTCATTTCTCCTCAGTACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 TCAAGTATGTCAGAGGATGCTGCTGCTTCTCAGACTGGGAGGCTGCTGAGTGTTCAGTT 960
DB 901 TCAAGTATGTCAGAGGATGCTGCTGCTTCTCAGACTGGGAGGCTGCTGAGTGTTCAGTT 960
QY 961 CCGTGTACAGCTAACCCAGTCAATTTGGACTCGAGAGGATATTTCTAATTAAGCTCCCC 1020
DB 961 CCGTGTACAGCTAACCCAGTCAATTTGGACTCGAGAGGATATTTCTAATTAAGCTCCCC 1020
QY 1021 AGAGCATCTCATAAAGCTGTGCTCAATGGAACATTTATGCTGGTGTGTTGGAGGATATG 1080
DB 1021 AGAGCATCTCATAAAGCTGTGCTCAATGGAACATTTATGCTGGTGTGTTGGAGGATATG 1080

QY 1081 TTCAACCACTCAGATTATAAANAGTGTCTAGCGTAGACCTTGCTCTAGGGAGTGGCTT 1140
DB 1081 TTCAACCACTCAGATTATAAANAGTGTCTAGCGTAGACCTTGCTCTAGGGAGTGGCTT 1140
QY 1141 CCACCTAAACCCGTCGTGTGGAACAATCGTGTGTAGATATGTCATCTTTTGGCATTATAC 1200
DB 1141 CCACCTAAACCCGTCGTGTGGAACAATCGTGTGTAGATATGTCATCTTTTGGCATTATAC 1200
QY 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260
DB 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260
QY 1261 TTGAGAGTTTTTCACTTCAATAGTCATGCGGTGTGTTGACCCCTAAGGCAAGAGAG 1320
DB 1261 TTGAGAGTTTTTCACTTCAATAGTCATGCGGTGTGTTGACCCCTAAGGCAAGAGAG 1320
QY 1321 CAGTATGCAGTGGTGGGCACTCTCTGACACACATTTTACACTGAAGAAATGGCCGAGTGGTC 1380
DB 1321 CAGTATGCAGTGGTGGGCACTCTCTGACACACATTTTACACTGAAGAAATGGCCGAGTGGTC 1380
QY 1381 ATGCTGTGTCATCTTTGGTCACTGCGCTCTCTATGGAATATTAAGCAATGTGACGAATAT 1440
DB 1381 ATGCTGTGTCATCTTTGGTCACTGCGCTCTCTATGGAATATTAAGCAATGTGACGAATAT 1440
QY 1441 GATTTGGATAGAACACATAGGATATTAACACCCAGGGTGCCCTTGTGCAAGGGGT 1500
DB 1441 GATTTGGATAGAACACATAGGATATTAACACCCAGGGTGCCCTTGTGCAAGGGGT 1500
QY 1501 TAGCGCCATPAGCAGTGTTTAGCAACATAGGACAGGGCCCTATAGTTTCATGTTGGTGTAC 1560
DB 1501 TAGCGCCATPAGCAGTGTTTAGCAACATAGGACAGGGCCCTATAGTTTCATGTTGGTGTAC 1560
QY 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGTGTGAGATGATCTCTACCGATATGATGTGGAT 1620
DB 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGTGTGAGATGATCTCTACCGATATGATGTGGAT 1620
QY 1621 ACCCAGATGTGGACCAATCTTAAGACACAGCCGATTTTCCGTTACTTGACACACAGCTGTG 1680
DB 1621 ACCCAGATGTGGACCAATCTTAAGACACAGCCGATTTTCCGTTACTTGACACACAGCTGTG 1680
QY 1681 ATAGTGTGAGTGNACATGCTGTGTGTTGGGGAAACACACAAATGACATCTATGAGC 1740
DB 1681 ATAGTGTGAGTGNACATGCTGTGTGTTGGGGAAACACACAAATGACATCTATGAGC 1740
QY 1741 CATGGCGCAATGCTCTCTTCAGATTTTCATGGCCTATGACATTTGCCCTGTGACCGCTGG 1800
DB 1741 CATGGCGCAATGCTCTCTTCAGATTTTCATGGCCTATGACATTTGCCCTGTGACCGCTGG 1800
QY 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAAAGATTTGGCCATTCAGCAGTC 1860
DB 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAAAGATTTGGCCATTCAGCAGTC 1860
QY 1861 TTACACAACAGCACCATGATGTGTTGGTGGTTCATAGTCTCTCTCAGCGACATC 1920
DB 1861 TTACACAACAGCACCATGATGTGTTGGTGGTTCATAGTCTCTCTCAGCGACATC 1920
QY 1921 CTGGTATTCACTCGGAACAGTGTGATGCGATCGGAGTGAAGCGCTGTTTATGACGCA 1980
DB 1921 CTGGTATTCACTCGGAACAGTGTGATGCGATCGGAGTGAAGCGCTGTTTATGACGCA 1980
QY 1981 GCACTGCTATTCGGTGTGTGGAACACAGGTCGTCAGTGTATCTCTGGCGCTG 2040
DB 1981 GCACTGCTATTCGGTGTGTGGAACACAGGTCGTCAGTGTATCTCTGGCGCTG 2040
QY 2041 GCAACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTTTGAC 2100
DB 2041 GCAACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTTTGAC 2100
QY 2101 CATGACAGATGTGACGACACACAGATTTGTATGAGCTGTACAGCCAAACCAATGACATGC 2160
DB 2101 CATGACAGATGTGACGACACACAGATTTGTATGAGCTGTACAGCCAAACCAATGACATGC 2160
QY 2161 CACTGGTGCATGACCAATGTGTGTCCTCCAGGAACACACAGCTGCTCAGAGGGCCAGATCTCC 2220

DB 2161 CACTGGTGCATGACCAATGTGTGTCCTCCAGGAACACACAGCTGCTCAGAGGGCCAGATCTCC 2220
QY 2221 ATTTTATAGGTATGAGAAATTCGCCCAAGGATAACCCCATGTACTACTGTATAAAGAGACCC 2280
DB 2221 ATTTTATAGGTATGAGAAATTCGCCCAAGGATAACCCCATGTACTACTGTATAAAGAGACCC 2280
QY 2281 AGCTGCAGAGAGCTGTGCCCTTGGACCCAGAACTGCGCAGTGGAGGCCCGGAATCAGGAGTGC 2340
DB 2281 AGCTGCAGAGAGCTGTGCCCTTGGACCCAGAACTGCGCAGTGGAGGCCCGGAATCAGGAGTGC 2340
QY 2341 ATTGCCCTGCCGAAAAATATCTGTGCAATTTGCTGTCATTTGTTGGAATCTCATGTTTG 2400
DB 2341 ATTGCCCTGCCGAAAAATATCTGTGCAATTTGCTGTCATTTGTTGGAATCTCATGTTTG 2400
QY 2401 AAAATTACTACTGCCAAGGAGAAATATGACAAATGCTATAATTTGTTCTGTAGGAACCAAT 2460
DB 2401 AAAATTACTACTGCCAAGGAGAAATATGACAAATGCTATAATTTGTTCTGTAGGAACCAAT 2460
QY 2461 GCCCTTTTGGCTTCTCTTCAACCCAGAGAGAGTGAATTTGCTCCTTAAGCAGCTGCGA 2520
DB 2461 GCCCTTTTGGCTTCTCTTCAACCCAGAGAGAGTGAATTTGCTCCTTAAGCAGCTGCGA 2520
QY 2521 ATAATGCAGTCACTCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGCGCCTTCGG 2580
DB 2521 ATAATGCAGTCACTCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGCGCCTTCGG 2580
QY 2581 AAGATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTTCAAAATAGTCTTACTA 2640
DB 2581 AAGATCAATGTCTACTGCTGCGAAGATATGTCCCAATTTTCAAAATAGTCTTACTA 2640
QY 2641 CAGTGAATGCGCTCTGAGCCAGTGTGCTGGAATTTCTGTGGAATTTTATCAGAACCCAGT 2700
DB 2641 CAGTGAATGCGCTCTGAGCCAGTGTGCTGGAATTTCTGTGGAATTTTATCAGAACCCAGT 2700
QY 2701 ACTCGGGAGCTGAAGGTGCAACCTGCATCAACCCACTCAATGGTGTGTCTGTAAAGG 2760
DB 2701 ACTCGGGAGCTGAAGGTGCAACCTGCATCAACCCACTCAATGGTGTGTCTGTAAAGG 2760
QY 2761 CCTGCAAAACACAGTGTAGCAGTGTCCGACACCATGTGCCCTTGAAGGACAGCATGTGGA 2820
DB 2761 CCTGCAAAACACAGTGTAGCAGTGTCCGACACCATGTGCCCTTGAAGGACAGCATGTGGA 2820
QY 2821 GATTGCACCAAGCGCAGCTCTGAGTGTGATGTGCGACCAACATGAAGCAGTGTGTGGAC 2880
DB 2821 GATTGCACCAAGCGCAGCTCTGAGTGTGATGTGCGACCAACATGAAGCAGTGTGTGGAC 2880
QY 2881 TCCATGCTTATGTGGCTCTCTCCCTTTTGGCAGTGTATGGAATGTTATACGATGAGC 2940
DB 2881 TCCATGCTTATGTGGCTCTCTCCCTTTTGGCAGTGTATGGAATGTTATACGATGAGC 2940
QY 2941 ACCTGCCCCCTGAAAAATTTTTCAGGCTACTGTACCTGTAGTCAATTTGCTTGAAGCAACCA 3000
DB 2941 ACCTGCCCCCTGAAAAATTTTTCAGGCTACTGTACCTGTAGTCAATTTGCTTGAAGCAACCA 3000
QY 3001 GGCTGTGGTGTGTACTGATCCAGCAATTTGCGCAAGGGGAAATGCAATAGAGGTTCC 3060
DB 3001 GGCTGTGGTGTGTACTGATCCAGCAATTTGCGCAAGGGGAAATGCAATAGAGGTTCC 3060
QY 3061 TATAAGGACCCAGTGAAGTGTCTTCCGACGCCCTTACAGGAATTTTCTATCCACAGCCC 3120
DB 3061 TATAAGGACCCAGTGAAGTGTCTTCCGACGCCCTTACAGGAATTTTCTATCCACAGCCC 3120
QY 3121 CTGCTCAATTTCCAGCATGTGTCTTAGAGGACAGCAGATACAACTGGTCTTTTCATTCAGTGT 3180
DB 3121 CTGCTCAATTTCCAGCATGTGTCTTAGAGGACAGCAGATACAACTGGTCTTTTCATTCAGTGT 3180
QY 3181 CCAAGCTTGCATGCAACCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240
DB 3181 CCAAGCTTGCATGCAACCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240
QY 3241 GAGAACCTGACCAACAGCAAGCACTGCGAGACCTGCAATCTCGCTTCTACCGTGTATCC 3300

Db 3241 GAGAACTGACACAGGCAAGCACTGCGAGACCTGATATCTGCTTCTACGGTGTATCCC 3300
Qy 3301 ACCAATGGAGGAAATGTGACGATCAAGTGGCAATGGGCAAGGCTCTCTGTGCAACACC 3360
Db 3301 ACCAATGGAGGAAATGTGACGATCAAGTGGCAATGGGCAAGGCTCTCTGTGCAACACC 3360
Qy 3361 AACACGGGCAAGTGTCTGACCAACCAAGGCGTCAAGGGGACGAGTGCAGCTATGT 3420
Db 3361 AACACGGGCAAGTGTCTGACCAACCAAGGCGTCAAGGGGACGAGTGCAGCTATGT 3420
Qy 3421 GAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTATCTCTCTT 3480
Db 3421 GAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTATCTCTCTT 3480
Qy 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATC 3540
Db 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATC 3540
Qy 3541 AATTTTGGTCTACTCTGACCAACAAACAGGATTTGGACATGTTCAATCAATGCCCTCC 3600
Db 3541 AATTTTGGTCTACTCTGACCAACAAACAGGATTTGGACATGTTCAATCAATGCCCTCC 3600
Qy 3601 AAGAAATTTCAACTCAACATCACTTGGGCTGGCTCTCAGCTGGAACCCAGGCTGGA 3660
Db 3601 AAGAAATTTCAACTCAACATCACTTGGGCTGGCTCTCAGCTGGAACCCAGGCTGGA 3660
Qy 3661 GAAGAGATGCTGTTGTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAAT 3720
Db 3661 GAAGAGATGCTGTTGTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAAT 3720
Qy 3721 GAGAGTTTGAATTTGCGAACCAACCAATATCACTTTCTTTGTTATGTCAAGTAATTC 3780
Db 3721 GAGAGTTTGAATTTGCGAACCAACCAATATCACTTTCTTTGTTATGTCAAGTAATTC 3780
Qy 3781 ACCTGCCCCATCAAAATTCAGATTCCTTCTCTACGACAGCAATTTTATGACCTGTA 3840
Db 3781 ACCTGCCCCATCAAAATTCAGATTCCTTCTCTACGACAGCAATTTTATGACCTGTA 3840
Qy 3841 CAGTTCTTCTGACATTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
Db 3841 CAGTTCTTCTGACATTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
Qy 3901 TCGAAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3960
Db 3901 TCGAAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3960
Qy 3961 CAACAGATGGCAGCGCTCTTGGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020
Db 3961 CAACAGATGGCAGCGCTCTTGGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020
Qy 4021 CTTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCAAACCCATTCGACTGGAGCGG 4080
Db 4021 CTTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCAAACCCATTCGACTGGAGCGG 4080
Qy 4081 TGTTTTGGCAACAAAGCCGCTCTCTCTGTTGTTGAGGCTCCCTCGAGGCGCTGGGT 4140
Db 4081 TGTTTTGGCAACAAAGCCGCTCTCTCTGTTGTTGAGGCTCCCTCGAGGCGCTGGGT 4140
Qy 4141 GGCATCCCTCTCTTGGGCGAGTCAGTCTTGTGTCGTCGAGCGCTCTGTCGACATTTCT 4200
Db 4141 GGCATCCCTCTCTTGGGCGAGTCAGTCTTGTGTCGTCGAGCGCTCTGTCGACATTTCT 4200
Qy 4201 CAGCAGATGCCGATAGTGTACAGGAGAGTACAGAGCCGTGAGAAACCGGAAGCAGCAG 4260
Db 4201 CAGCAGATGCCGATAGTGTACAGGAGAGTACAGAGCCGTGAGAAACCGGAAGCAGCAG 4260
Qy 4261 CCCCTGACACGCTGGGACCTGCAATCTGA 4290
Db 4261 CCCCTGACACGCTGGGACCTGCAATCTGA 4290

ID AAS72659 standard; cDNA; 8561 BP.
XX
AC AAS72659;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8463.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08472.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8463; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Query Match 100.0%; Score 4290; DB 5; Length 8561;
Best Local Similarity 100.0%; Pred No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGTGGCCGCGACGCGCGCAACTGAGCAAGGCTGAGGAGGAGACGCGCGGCGGCA 60
Db 1 ATGGTGGCCGCGACGCGCGCAACTGAGCAAGGCTGAGGAGGAGACGCGCGGCGGCA 60
Qy 61 GCGCTCGCGGCGAGGAGCGCGCGGCTGGGACTGGGCTGACACGAGGCTGGGAGG 120
Db 61 GCGCTCGCGGCGAGGAGCGCGCGGCTGGGACTGGGCTGACACGAGGCTGGGAGG 120
Qy 121 CCGGGGCTGGGGCGCGGGCTGGGCTCCCGCGGCTGCTCTCCACCGCTGGCGGCGACGG 180

2341 ATTTGCTCCCGAAAAATATCTGTGGCATTTGGTGGAAATCATGTTTG 2400
Db
2341 ATTTGCTCCCGAAAAATATCTGTGGCATTTGGTGGAAATCATGTTTG 2400
Qy
2401 AAAATTTACTTGCAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACACAAAT 2460
Db
2401 AAAATTTACTTGCAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACACAAAT 2460
Qy
2461 GGCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTGGGA 2520
Db
2461 GGCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTGGGA 2520
Qy
2521 ATAAATGCAATCTATCAGAGCATGTCAGAGCTCACTTAAACCCATGGGTGGCTTCGG 2580
Db
2521 ATAAATGCAATCTATCAGAGCATGTCAGAGCTCACTTAAACCCATGGGTGGCTTCGG 2580
Qy
2581 AAGATCAATGCTCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGTTTACTA 2640
Db
2581 AAGATCAATGCTCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGTTTACTA 2640
Qy
2641 CAGTGATGCCCTCTGAGCCAGTGCATGCTGGAATTTCTGTGAATTTTATCAAAACCCAGT 2700
Db
2641 CAGTGATGCCCTCTGAGCCAGTGCATGCTGGAATTTCTGTGAATTTTATCAAAACCCAGT 2700
Qy
2701 ACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAATGCTGCTGGAAG 2760
Db
2701 ACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAATGCTGCTGGAAG 2760
Qy
2761 CTTGCAAAACCAAGTCTAAGCAGTGCAGGACCACTGCTGAGGACAGCATGTGGA 2820
Db
2761 CTTGCAAAACCAAGTCTAAGCAGTGCAGGACCACTGCTGAGGACAGCATGTGGA 2820
Qy
2821 GATTGACACAGGCGAGCTCTGAGTGCAATGCTGCTGAGGACCACTGAGCAGTGTGAGAC 2880
Db
2821 GATTGACACAGGCGAGCTCTGAGTGCAATGCTGCTGAGGACCACTGAGCAGTGTGAGAC 2880
Qy
2881 TCCAAATGCTATGTGCTCTCTTCCCTTTGCGCAGTGTATGGAATGATACGATGAGC 2940
Db
2881 TCCAAATGCTATGTGCTCTCTTCCCTTTGCGCAGTGTATGGAATGATACGATGAGC 2940
Qy
2941 ACCTGCCCCCTGAAAAATTTGTTGAGGCTACTGATCTGATCTGATCTGCTGAGCAACCA 3000
Db
2941 ACCTGCCCCCTGAAAAATTTGTTGAGGCTACTGATCTGATCTGATCTGCTGAGCAACCA 3000
Qy
3001 GCTGTGGCTGTGCTACTGATCCAGCAATCTGCGAAGGCAATGATAGAGGTTCC 3060
Db
3001 GCTGTGGCTGTGCTACTGATCCAGCAATCTGCGAAGGCAATGATAGAGGTTCC 3060
Qy
3061 TATAAAGGACCAAGTGAAGATGCTTTCGAAAGCCCTACAGGAAATTTCTATCCAGGCC 3120
Db
3061 TATAAAGGACCAAGTGAAGATGCTTTCGAAAGCCCTACAGGAAATTTCTATCCAGGCC 3120
Qy
3121 CTGCTCAATTCAGCATGTGCTAGAGGACAGCATACAACTGGTCTTTCACTCACTGT 3180
Db
3121 CTGCTCAATTCAGCATGTGCTAGAGGACAGCATACAACTGGTCTTTCACTCACTGT 3180
Qy
3181 CCAGCTTGCAATGCAACCGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGT 3240
Db
3181 CCAGCTTGCAATGCAACCGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGT 3240
Qy
3241 GAGAACCTTGACCAAGGCAAGCATCTGCGAGACCTGCATATCTGGCTTCTACGGTATCCC 3300
Db
3241 GAGAACCTTGACCAAGGCAAGCATCTGCGAGACCTGCATATCTGGCTTCTACGGTATCCC 3300
Qy
3301 ACCAATGAGGGAATGTGACGATGCAAGTGCATGCAATGCGCTCTCTGTGCAACACC 3360
Db
3301 ACCAATGAGGGAATGTGACGATGCAAGTGCATGCAATGCGCTCTCTGTGCAACACC 3360
Qy
3361 AACACGGGCAAGTCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCACCTATGT 3420
Db
3361 AACACGGGCAAGTCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCACCTATGT 3420

3421 GAGGTAGAAAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTTATCTCTTCTT 3480
Db
3421 GAGGTAGAAAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTTATCTCTTCTT 3480
Qy
3481 ATTGACTATCAGTTCACTTCTAGTCTATCCAGGAGATGATCGTATTATACAGCTATC 3540
Db
3481 ATTGACTATCAGTTCACTTCTAGTCTATCCAGGAGATGATCGTATTATACAGCTATC 3540
Qy
3541 AATTTTGTGGCTACTCTCTGACGAAACAAAACAGGGATTTGGACATGTTCTCAATCCCTCC 3600
Db
3541 AATTTTGTGGCTACTCTCTGACGAAACAAAACAGGGATTTGGACATGTTCTCAATCCCTCC 3600
Qy
3601 AAGAAATTTCAACCTCAACATCACCCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGA 3660
Db
3601 AAGAAATTTCAACCTCAACATCACCCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGA 3660
Qy
3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGATACAAAGATAGTTTCTCTAAT 3720
Db
3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGATACAAAGATAGTTTCTCTAAT 3720
Qy
3721 GAGAAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTATGTCAGTAATTC 3780
Db
3721 GAGAAATTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTATGTCAGTAATTC 3780
Qy
3781 ACCTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAAAATTTATGACCTGTA 3840
Db
3781 ACCTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAAAATTTATGACCTGTA 3840
Qy
3841 CAGTTCTCTGATCTTTTCAAGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
Db
3841 CAGTTCTCTGATCTTTTCAAGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
Qy
3901 TGAAGATCAAAACAAAGTTTGGGCTCCAGACCTAGAGAGCACTTCTTCGAGAGATG 3960
Db
3901 TGAAGATCAAAACAAAGTTTGGGCTCCAGACCTAGAGAGCACTTCTTCGAGAGATG 3960
Qy
3961 CAACAGATGCCAGCGCTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020
Db
3961 CAACAGATGCCAGCGCTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020
Qy
4021 CCTCTGATCTTATTGGGGAGATATAAGACTGTTCCTCAAAACCCATTCGACTGGAGCG 4080
Db
4021 CCTCTGATCTTATTGGGGAGATATAAGACTGTTCCTCAAAACCCATTCGACTGGAGCG 4080
Qy
4081 TGTCTTGGCAACAAAGCCGCT 4140
Db
4081 TGTCTTGGCAACAAAGCCGCT 4140
Qy
4141 GGCATCCCTCTCTCTGGGAGTCAAGTCTTGTCTGTGGCCAGCGCTCTGTGTGACATTTCT 4200
Db
4141 GGCATCCCTCTCTCTGGGAGTCAAGTCTTGTCTGTGGCCAGCGCTCTGTGTGACATTTCT 4200
Qy
4201 CAGCAGATCCGATAGTGTACAGGAGAGTCAAGAGCGCTGAGAAACCGGAGAGCAGAG 4260
Db
4201 CAGCAGATCCGATAGTGTACAGGAGAGTCAAGAGCGCTGAGAAACCGGAGAGCAGAG 4260
Qy
4261 CCCCCCTGCAAGCCTGGGACCTGCATCTGA 4290
Db
4261 CCCCCCTGCAAGCCTGGGACCTGCATCTGA 4290

RESULT 3

AAZ52277

ID AAZ52277 standard; cDNA; 3819 BP.

XX AAZ52277;

XX AC AC

DT 18-JUL-2000 (first entry)

XX Human soluble attractin-2 cDNA.

XX Human; soluble attractin-2; immune response; macrophage; monocyte;

KW T cell; immunostimulant; immunosuppressed patient; cancer;

immunodeficiency syndrome; transplant; autoimmune disease; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 1..3819

FT /*tag= a

FT /product= "Soluble attractin-2"

PN WO200015651-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-USO20948.

XX 14-SEP-1998; 98US-0100137P.

PA (DAND) DANA FARBER CANCER INST INC.

PI Duke-Cohan JS, Schlosman SF;

XX WPI; 2000-271373/23.

DR P-PSDB; AAY70692.

XX Isolated nucleic acids encoding human attractin polypeptides useful for enhancing immune responses.

XX Claim 3; Fig 12; 120pp; English.

XX The patent discloses four forms of human attractin polypeptides which enhance immune response by promoting macrophage and monocyte spreading in the presence of T cells. These include soluble attractin-1 and -2 and membrane attractin-1 and -2. These various forms of attractin are encoded by alternatively spliced mRNA molecule transcribed from a single gene. The present sequence is a cDNA encoding human soluble attractin-2.

XX Attractin-2 differs from attractin-1 in having a 74 amino acid insert in the N-terminal. This sequence can be used to enhance immune response in immunosuppressed patients such as those undergoing chemo- and radio-therapy treatment for cancer or those suffering from common variable immunodeficiency syndrome. The proteins may also be used to screen modulators (agonists and antagonists) of immune response which may also be used to regulate immune reactions. Attractin antibodies can be used to inhibit immune response in transplant recipients or patients afflicted with autoimmune disease

XX SQ Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

Query Match 88.6%; Score 3801; DB 3; Length 3819;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGGCCGAGCGCGCACTGAGCAAGGCTGAGGAGGAGCGCGGCGACGGCA 60

DB 1 ATGGTGGCCGAGCGCGCACTGAGCAAGGCTGAGGAGGAGCGCGGCGACGGCA 60

QY 61 CGCTTCGGCGCAGAGCGCGCGCGCACTGGGACTGGGACGTGGACAGCGGTGGGAGG 120

DB 61 CGCTTCGGCGCAGAGCGCGCGCGCACTGGGACTGGGACGTGGACAGCGGTGGGAGG 120

QY 121 CCGGGCTGGGGCGCGGCTCGCGCTCCCGGGCTGTCTCCACCGCTGCGGCGCACGG 180

DB 121 CCGGGCTGGGGCGCGGCTCGCGCTCCCGGGCTGTCTCCACCGCTGCGGCGCACGG 180

QY 181 CTGCTGCTGCTGCTGCTGCTGCTCCCGCGCGCTGTGCTGCTGCTGCTGCTGCTGAG 240

DB 181 CTGCTGCTGCTGCTGCTGCTGCTCCCGCGCGCTGTGCTGCTGCTGCTGCTGCTGAG 240

QY 241 GCCGAGCGCGCGCGCGCGCGCGCTGCGGGCTCAGCGCAGCGCGCGCGCAAGGAA 300

DB 241 GCCGAGCGCGCGCGCGCGCGCGCTGCGGGCTCAGCGCAGCGCGCGCGCAAGGAA 300

QY 301 TGTGACCGCGCTGTGTCAACGGCGGCTCGCTGCACCGCTGGCACCGGCGCGCTGTCG 360

DB 301 TGTGACCGCGCTGTGTCAACGGCGGCTCGCTGCACCGCTGGCACCGGCGCGCTGTCG 360

QY 361 CCCGCCGCTGGGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCCAGACTACTGGA 420

DB 361 CCCGCCGCTGGGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCCAGACTACTGGA 420

QY 421 TCTTCTGGGTTTGTGACAGATGGACCTGGAATATAAATACAAACCAAGTGCACGTGG 480

DB 421 TCTTCTGGGTTTGTGACAGATGGACCTGGAATATAAATACAAACCAAGTGCACGTGG 480

QY 481 CTCAATTGAAGGACAGCCAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAG 540

DB 481 CTCAATTGAAGGACAGCCAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAG 540

QY 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGT 600

DB 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGT 600

QY 601 GCATTTAGTGGCTCATTTGTTTCTCGAGAGAGATGCAATGAGACTGTCCTGAGGTTGT 660

DB 601 GCATTTAGTGGCTCATTTGTTTCTCGAGAGAGATGCAATGAGACTGTCCTGAGGTTGT 660

QY 661 GCCACATCAGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGA 720

DB 661 GCCACATCAGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGA 720

QY 721 TTTAATATTACTACAGTTTTCATATGTTCCAAATAAATGCTCAGGCGGAGAGAGTGT 780

DB 721 TTTAATATTACTACAGTTTTCATATGTTCCAAATAAATGCTCAGGCGGAGAGAGTGT 780

QY 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTAATGTTCTGAAACCTGGAAGGTGAA 840

DB 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTAATGTTCTGAAACCTGGAAGGTGAA 840

QY 841 GCATGTGACATTCCTCAGTACAGCACTGCTGTTTCTCATCGAGGCAATCTGCAAT 900

DB 841 GCATGTGACATTCCTCAGTACAGCACTGCTGTTTCTCATCGAGGCAATCTGCAAT 900

QY 901 TCAAGTGTGTGAGAGGATGCTCCTGCTTCTCAGACTGCGAGGGTCTGGATGTTCAATT 960

DB 901 TCAAGTGTGTGAGAGGATGCTCCTGCTTCTCAGACTGCGAGGGTCTGGATGTTCAATT 960

QY 961 CTTGTACGAGCTAACGAGTCATTTGGACTCAGAGGAATATTTCAACTTAAGTCCCG 1020

DB 961 CTTGTACGAGCTAACGAGTCATTTGGACTCAGAGGAATATTTCAACTTAAGTCCCG 1020

QY 1021 AGAGCATCTCAATAAGCTGTGCTCAATGAAACATTTATGTTGGTGTGGAGATATATG 1080

DB 1021 AGAGCATCTCAATAAGCTGTGCTCAATGAAACATTTATGTTGGTGTGGAGATATATG 1080

QY 1081 TTCAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTT 1140

DB 1081 TTCAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTT 1140

QY 1141 CCACATAACCGTTCTGTGAACAATGTTGTTGTAGATATGGTCTTTTGGCATTTATAC 1200

DB 1141 CCACATAACCGTTCTGTGAACAATGTTGTTGTAGATATGGTCTTTTGGCATTTATAC 1200

QY 1201 AAGGATAAATTTACATGTATGAGGAAATTTGATCCAACTGGGAATGTGCCAATGAG 1260

DB 1201 AAGGATAAATTTACATGTATGAGGAAATTTGATCCAACTGGGAATGTGCCAATGAG 1260

QY 1261 TTGAGAGTTTTCACATTCATTAATGAGTCAATGGGTGTTGTTGACCCCTAAGCAAGGAG 1320

DB 1261 TTGAGAGTTTTCACATTCATTAATGAGTCAATGGGTGTTGTTGACCCCTAAGCAAGGAG 1320

QY 1321 CAGTATGAGTGGTTGGGCACTCTGCACACATTTGTACACTGAAGAATGGCGAGTGGTC 1380

DB 1321 CAGTATGAGTGGTTGGGCACTCTGCACACATTTGTACACTGAAGAATGGCGAGTGGTC 1380

QY 1381 ATGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGATAT 1440

DB 1381 ATGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGATAT 1440

1441 GATTTGGATAAGAACACATATGGAGTATATATACACACCCAGGGTGCCTTTGTGCAAGGGGGT 1500
1441 GATTTGGATAAGAACACATATGGAGTATATATACACACCCAGGGTGCCTTTGTGCAAGGGGGT 1500
1501 TAGGGCCATAGCAGTGTTCAGACCATAGGACCGGGCCCTATACGTTTCATGGTGGCTAC 1560
1501 TAGGGCCATAGCAGTGTTCAGACCATAGGACCGGGCCCTATACGTTTCATGGTGGCTAC 1560
1561 AAGGCTTTTCAGTGCCTTAAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGAT 1620
1561 AAGGCTTTTCAGTGCCTTAAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGAT 1620
1621 ACCGAGATGAGCAATCTTTAAGACAGCGGATTTTCCGGTATCTTGCAACAGCTGTG 1680
1621 ACCGAGATGAGCAATCTTTAAGACAGCGGATTTTCCGGTATCTTGCAACAGCTGTG 1680
1681 ATAGTGAAGTGAACCATGCTGGTGTTCGGGGAAACACACACATGACATCTATGAGC 1740
1681 ATAGTGAAGTGAACCATGCTGGTGTTCGGGGAAACACACACATGACATCTATGAGC 1740
1741 CATGGCCCAATGCTTCTTCATGATTTTCATGCTTATGCTTATGCAATGCTGTGACCGGTG 1800
1741 CATGGCCCAATGCTTCTTCATGATTTTCATGCTTATGCTTATGCAATGCTGTGACCGGTG 1800
1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTC 1860
1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTC 1860
1861 TTACACACAGCACCATGATGTTGGTGGTTCATAGTCTCTCCCTCAGCGACATC 1920
1861 TTACACACAGCACCATGATGTTGGTGGTTCATAGTCTCTCCCTCAGCGACATC 1920
1921 CTGGTATTCCTCGGACAGTGTATGCTGATCGGATGAGGCGCTTGTATTAGCAGCA 1980
1921 CTGGTATTCCTCGGACAGTGTATGCTGATCGGATGAGGCGCTTGTATTAGCAGCA 1980
1981 GGACCTGTATTCGGTGTGTGGAACACAGGGTCTGCTCAGTGTATCTCTGGGGCGCTG 2040
1981 GGACCTGTATTCGGTGTGTGGAACACAGGGTCTGCTCAGTGTATCTCTGGGGCGCTG 2040
2041 GCACTGTATGAACAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100
2041 GCACTGTATGAACAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100
2101 CATGACAGATGTGACGACACACAGATTTGTACAGCTGTACAGCAACACCAATGACTGC 2160
2101 CATGACAGATGTGACGACACACAGATTTGTACAGCTGTACAGCAACACCAATGACTGC 2160
2161 CACTGGTGCATGACCATTTGTGTCGCCAGGAAACACAGCTGCTCAGAAAGCCAGATCTCC 2220
2161 CACTGGTGCATGACCATTTGTGTCGCCAGGAAACACAGCTGCTCAGAAAGCCAGATCTCC 2220
2221 ATTTTGTAGTATGAATTCGCCAGAGTAAACCCATGCTACTGTACAGAGACAC 2280
2221 ATTTTGTAGTATGAATTCGCCAGAGTAAACCCATGCTACTGTACAGAGACAC 2280
2281 AGCTGCAGGAGCTGTGCCCTTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGC 2340
2281 AGCTGCAGGAGCTGTGCCCTTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGC 2340
2341 ATTGGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAAACTCATGTTTG 2400
2341 ATTGGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAAACTCATGTTTG 2400
2401 AAAATTTACTACTGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAAT 2460
2401 AAAATTTACTACTGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAAT 2460
2461 GCCCTTTGGCTTCTCTTAACCCAGAGAGAGTAAATTTGCTCTTAAGAGAGCTGCGGA 2520
2461 GCCCTTTGGCTTCTCTTAACCCAGAGAGAGTAAATTTGCTCTTAAGAGAGCTGCGGA 2520

2521 ATAAATCAGTCTATCTCAGAGCATGTCCAAAGCTACCTTAAACCCCATGGGTGCGCTTCGG 2580
2521 ATAAATCAGTCTATCTCAGAGCATGTCCAAAGCTACCTTAAACCCCATGGGTGCGCTTCGG 2580
2581 AAGATCAATGTGTCTTACTGGTGTGGGAGATATGTCCCATTTTACAAATAGTTTACTA 2640
2581 AAGATCAATGTGTCTTACTGGTGTGGGAGATATGTCCCATTTTACAAATAGTTTACTA 2640
2641 CAGTGTATCCCGTCTTGAGCCCAAGTGTGTAATCTGTGGAAATTTTATCAGAACCCAGT 2700
2641 CAGTGTATCCCGTCTTGAGCCCAAGTGTGTAATCTGTGGAAATTTTATCAGAACCCAGT 2700
2701 ACTCGGGGACTGAGCTGCAACCTGCACTCAACCCCACTCAATGGTAGTGTCTGTGAAAGG 2760
2701 ACTCGGGGACTGAGCTGCAACCTGCACTCAACCCCACTCAATGGTAGTGTCTGTGAAAGG 2760
2761 CCTGCAAAACACAGTCTTAAGCAGTGCAGACACCATGTGCTTGGAGCAGCATGTGGA 2820
2761 CCTGCAAAACACAGTCTTAAGCAGTGCAGACACCATGTGCTTGGAGCAGCATGTGGA 2820
2821 GATTCACACAGCGGAGCTCTGAGTGCATGTGCGGACACCATGTGCTTGGAGCAGCATGTGGA 2880
2821 GATTCACACAGCGGAGCTCTGAGTGCATGTGCGGACACCATGTGCTTGGAGCAGCATGTGGA 2880
2881 TCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTTATACAGATGAGC 2940
2881 TCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTTATACAGATGAGC 2940
2941 ACCTGCCCCCTGAAAATTTGTTTCCAGCTACTGTACTCTGTAGTCAATTTGTTGGAGCAACA 3000
2941 ACCTGCCCCCTGAAAATTTGTTTCCAGCTACTGTACTCTGTAGTCAATTTGTTGGAGCAACA 3000
3001 GCGTGTGGTGTGTACTGATCCAGCAATACTGCGCAAGGGAATGATAGAGGTTCC 3060
3001 GCGTGTGGTGTGTACTGATCCAGCAATACTGCGCAAGGGAATGATAGAGGTTCC 3060
3061 TATAAGGACCAAGTGAAGTCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGGCC 3120
3061 TATAAGGACCAAGTGAAGTCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGGCC 3120
3121 CTGCTCAATTCAGCATGTGTCTAGAGGACAGCATACACTGCTCTTCTTCACTGT 3180
3121 CTGCTCAATTCAGCATGTGTCTAGAGGACAGCATACACTGCTCTTCTTCACTGT 3180
3181 CCAGCTTGCATATGCAAGCGGCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGT 3240
3181 CCAGCTTGCATATGCAAGCGGCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGT 3240
3241 GAGAACCTTGACACAGGCAAGCACTTGCAGACCTGCTCATATCTGGCTTCTAGCGTATCCC 3300
3241 GAGAACCTTGACACAGGCAAGCACTTGCAGACCTGCTCATATCTGGCTTCTAGCGTATCCC 3300
3301 ACCAATGGAGGAAATGTGAGCATGCAAGTGCATGCGGACCGCTCTCTGTGCAACACC 3360
3301 ACCAATGGAGGAAATGTGAGCATGCAAGTGCATGCGGACCGCTCTCTGTGCAACACC 3360
3361 AACACGGGCAAGTGTCTTGACACCAAGGGCGTCAAGGGGGAGAGTGCAGCTATGT 3420
3361 AACACGGGCAAGTGTCTTGACACCAAGGGCGTCAAGGGGGAGAGTGCAGCTATGT 3420
3421 GAGGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTT 3480
3421 GAGGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTT 3480
3481 ATTGACTATCAGTTTCACTATCCAGGAAGATGATCGCTATTAACAGCTATC 3540
3481 ATTGACTATCAGTTTCACTATCCAGGAAGATGATCGCTATTAACAGCTATC 3540
3541 AATTTTGGCTTCTCTTGAACAAAGGAGTTTGGACATGTTTCAATGCTCTCC 3600
3541 AATTTTGGCTTCTCTTGAACAAAGGAGTTTGGACATGTTTCAATGCTCTCC 3600
3601 AAGAATTTCAACCTCAACATCAGTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660

Db 3601 AAGAATTCAACCTCAACATCACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660
Qy 3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGGTACAAAGATAGTTTCTCTAAT 3720
Db 3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGGTACAAAGATAGTTTCTCTAAT 3720
Qy 3721 GAGAAGTTTGATTTGCAACCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTC 3780
Db 3721 GAGAAGTTTGATTTGCAACCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTC 3780
Qy 3781 ACCTGGCCCATCAAAATTCAG 3801
Db 3781 ACCTGGCCCATCAAAATTCAG 3801
RESULT 4
AAZ52275
ID AAZ52275 standard; cDNA; 4068 BP.
XX AC
XX AAZ52275;
XX 18-JUL-2000 (first entry)
XX Human membrane attractin-1 cDNA.
XX Human; membrane attractin-1; immune response; macrophage; monocyte;
XX T cell; immunostimulant; immunosuppressed patient; cancer;
XX immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..4068
XX /*tag= a
XX /product= "Membrane attractin-1"
XX WO200015651-A1.
XX 23-MAR-2000.
XX 14-SEP-1999; 99WO-US020948.
XX 14-SEP-1998; 98US-0100137P.
XX (DAND) DANA FARBER CANCER INST INC.
XX Duke-Cohan JS, Schlossman SF;
XX WPI; 2000-271373/23.
XX P-PSDB; RAY70690.
XX Isolated nucleic acids encoding human attractin polypeptides useful for
XX enhancing immune responses.
XX Claim 3; Fig 10; 120pp; English.
XX The patent discloses four forms of human attractin polypeptides which
XX enhance immune response by promoting macrophage and monocyte spreading in
XX the presence of T cells. These include soluble attractin-1 and -2 and
XX membrane attractin-1 and -2. These various forms of attractin are encoded
XX by alternatively spliced mRNA molecule transcribed from a single gene.
XX The present sequence is a cDNA encoding human membrane attractin-1.
XX Membrane attractin differs from soluble attractin in having a
XX transmembrane domain and a cytoplasmic domain at the C-terminal. This
XX sequence can be used to enhance immune response in immunosuppressed
XX patients such as those undergoing chemo- and radio-therapy treatment for
XX cancer or those suffering from common variable immunodeficiency syndrome.
XX The proteins may also be used to screen modulators (agonists and
XX antagonists) of immune responses which may also be used to regulate immune
XX reactions. Attractin antibodies can be used to inhibit immune response in
XX transplant recipients or patients afflicted with autoimmune disease

SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;
Query Match 88.0%; Score 3775; DB 3; Length 4068;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 3975; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 312 CTGTGTCACACGGCGTCCCTGCACACCCCTGCACCGCCAGTGCCTCCCGCGCGGTG 371
Db 90 CTGTGTCACACGGCGTCCCTGCACACCCCTGCACCGCCAGTGCCTCCCGCGCGGTG 149
Qy 372 GGTGGGCGAGCAATCCAGACACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTT 431
Db 150 GGTGGGCGAGCAATCCAGACACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTT 209
Qy 432 TGTGACAGATGACCTCGAAATTAATAACAAACAGAGTGCACGTGGCTCATTAAGAG 491
Db 210 TGTGACAGATGACCTCGAAATTAATAACAAACAGAGTGCACGTGGCTCATTAAGAG 269
Qy 492 ACAGCAAAATAGAATAATGAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGGA 551
Db 270 ACAGCAAAATAGAATAATGAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGGA 329
Qy 552 CCATTTATATGTTTATGATGGGACTCAATTTATGACCGGCTAGTTGCTGATTTAGTGG 611
Db 330 CCATTTATATGTTTATGATGGGACTCAATTTATGACCGGCTAGTTGCTGATTTAGTGG 389
Qy 612 CCTCATTTGTTCTGAGAGAGATGGCAATGAGACTCTCCCTGAGGTTGTTGCCACATCAGG 671
Db 390 CCTCATTTGTTCTGAGAGAGATGGCAATGAGACTCTCCCTGAGGTTGTTGCCACATCAGG 449
Qy 672 TTATGCTTGTGTCATTTTTTTTAGTGATGCTGCTTATAATTTAGCTGGATTTAATTTAC 731
Db 450 TTATGCTTGTGTCATTTTTTTTAGTGATGCTGCTTATAATTTAGCTGGATTTAATTTAC 509
Qy 732 TTACAGTTTGTATGTTGTCGAAATTAATGCTCAGCGCGAGGAGAGTGTAGATCAGTAA 791
Db 510 TTACAGTTTGTATGTTGTCGAAATTAATGCTCAGCGCGAGGAGAGTGTAGATCAGTAA 569
Qy 792 TAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACTGGAAGTGAAGCATGTGACAT 851
Db 570 TAGCAGCGATACTGTTGAATGTGAATGTTCTGAAACTGGAAGTGAAGCATGTGACAT 629
Qy 852 TCCTCAGCTACAGACAACCTGTTGTTTCTCATCGAGGCATCTGCAATTCAGTGTGT 911
Db 630 TCCTCAGCTACAGACAACCTGTTGTTTCTCATCGAGGCATCTGCAATTCAGTGTGT 689
Qy 912 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTGATTCCTGTACCAGC 971
Db 690 CAGAGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTGATTCCTGTACCAGC 749
Qy 972 TAACCACTCATTTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAGAGCATCTCA 1031
Db 750 TAACCACTCATTTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAGAGCATCTCA 809
Qy 1032 TAAAGCTGTGTCATGGAACATTTATGTTGGGTTGTTGGAGGATATATGTTCAACCACTC 1091
Db 810 TAAAGCTGTGTCATGGAACATTTATGTTGGGTTGTTGGAGGATATATGTTCAACCACTC 869
Qy 1092 AGATTATACATGTTCTAGCGTATGACCTTCTTCTAGGGAGTGGCTTCCACTAAACCG 1151
Db 870 AGATTATACATGTTCTAGCGTATGACCTTCTTCTAGGGAGTGGCTTCCACTAAACCG 929
Qy 1152 TTCTGTGAACAATGTTGTTGTTAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1211
Db 930 TTCTGTGAACAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 989
Qy 1212 TTACATGTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1271
Db 990 TTACATGTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1049
Qy 1272 TCACATTCATATGAGTCAATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1331
Db 1050 TCACATTCATATGAGTCAATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1109

QY 1332 GGTGGGCACTCTGCACACATGTTTACACTGAAGAATGGCCGAGTGGTCAATGCTGGTCAAT 1391
DB 1110 GGTGGGCACTCTGCACACATGTTTACACTGAAGAATGGCCGAGTGGTCAATGCTGGTCAAT 1169
QY 1392 CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAA 1451
DB 1170 CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAA 1229
QY 1452 GAACATGAGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAG 1511
DB 1230 GAACATGAGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAG 1289
QY 1512 CAGTGTATACGACCATAGGACCGGCCCTATAGTTCATGGTGGCTACAGGCTTTTCAG 1571
DB 1290 CAGTGTATACGACCATAGGACCGGCCCTATAGTTCATGGTGGCTACAGGCTTTTCAG 1349
QY 1572 TGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG 1631
DB 1350 TGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG 1409
QY 1632 GACCATTTTAAGGACACCGCATTTTCCGTTACTTGCACACACTGTGATAGTGGTGG 1691
DB 1410 GACCATTTTAAGGACACCGCATTTTCCGTTACTTGCACACACTGTGATAGTGGTGG 1469
QY 1692 AACCATGCTGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCAA 1751
DB 1470 AACCATGCTGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCAA 1529
QY 1752 ATGCTTCTCTTACAGATTCATGGGCTATGACATTCCTGTGACCGCTGGTCAATGCTTCC 1811
DB 1530 ATGCTTCTCTTACAGATTCATGGGCTATGACATTCCTGTGACCGCTGGTCAATGCTTCC 1589
QY 1812 CAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGCTTTACACAACAG 1871
DB 1590 CAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGCTTTACACAACAG 1649
QY 1872 CACCATGATGTTTGGTGGTTCATATGATCTCTCCTCAGGACATCTCTGATTTAC 1931
DB 1650 CACCATGATGTTTGGTGGTTCATATGATCTCTCCTCAGGACATCTCTGATTTAC 1709
QY 1932 CTCGGAACAGTGCATCGGATCGGATGAGCGCTTGTAGCAGCAGGACCTGGTAT 1991
DB 1710 CTCGGAACAGTGCATCGGATCGGATGAGCGCTTGTAGCAGCAGGACCTGGTAT 1769
QY 1992 TCGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGA 2051
DB 1770 TCGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGA 1829
QY 2052 ACAGAGAAAAGTTAAATCAGAAATGTTTTCCAAAAGAACTCTTGAACATGACAGATG 2111
DB 1830 ACAGAGAAAAGTTAAATCAGAAATGTTTTCCAAAAGAACTCTTGAACATGACAGATG 1889
QY 2112 TGACCAACACACAGATTTTACAGCTGTACAGCCACACCAATGACTGGCCACTGGTGCAA 2171
DB 1890 TGACCAACACACAGATTTTACAGCTGTACAGCCACACCAATGACTGGCCACTGGTGCAA 1949
QY 2172 TGACCAATGTTTCCCAAGAAACACACAGCTGTACAGAGGCGCAGATCTCCATTTTAGTTA 2231
DB 1950 TGACCAATGTTTCCCAAGAAACACACAGCTGTACAGAGGCGCAGATCTCCATTTTAGTTA 2009
QY 2232 TGAGAAATGGCCCAAGGATTAACCCATGACTACTGTAAACAGAACACAGCTGCGAGGAG 2291
DB 2010 TGAGAAATGGCCCAAGGATTAACCCATGACTACTGTAAACAGAACACAGCTGCGAGGAG 2069
QY 2292 CTGTGCCCTGGAACAGAACTGCCAGTGGGAGCCCCCGAATCAGGAGTGCATTGCCCTGCC 2351
DB 2070 CTGTGCCCTGGAACAGAACTGCCAGTGGGAGCCCCCGAATCAGGAGTGCATTGCCCTGCC 2129
QY 2352 CGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTGAAATTTACTAC 2411
DB 2130 CGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTGAAATTTACTAC 2189

QY 2412 TGCCAAGGAGAAATTAAGCAATGCTAAATGTTCTGTAGGAACCAATGCCCTTTTGGC 2471
DB 2190 TGCCAAGGAGAAATTAAGCAATGCTAAATGTTCTGTAGGAACCAATGCCCTTTTGGC 2249
QY 2472 TTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTGCGAAATAAGCAGTC 2531
DB 2250 TTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTGCGAAATAAGCAGTC 2309
QY 2532 ATCTCAGAGCATGTCAAAGCTCACCTTAAACCCATGGGTGGCCCTTCGGAAGATCAATGT 2591
DB 2310 ATCTCAGAGCATGTCAAAGCTCACCTTAAACCCATGGGTGGCCCTTCGGAAGATCAATGT 2369
QY 2592 GTCTACTGCTGCTGGAGAGATATGTCGCCCATTTTACAAATAGTTTACTACAGTGGATGCC 2651
DB 2370 GTCTACTGCTGCTGGAGAGATATGTCGCCCATTTTACAAATAGTTTACTACAGTGGATGCC 2429
QY 2652 GTCTGAGCCAGTGTGATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACT 2711
DB 2430 GTCTGAGCCAGTGTGATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACT 2489
QY 2712 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGNAAGCCCTGCAAAACA 2771
DB 2490 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGNAAGCCCTGCAAAACA 2549
QY 2772 CAGTGTCTAAGCAGTGTCCGACACCATGTGCTTGAAGCAGCATGTGAGAGATTGCACCAG 2831
DB 2550 CAGTGTCTAAGCAGTGTCCGACACCATGTGCTTGAAGCAGCATGTGAGAGATTGCACCAG 2609
QY 2832 CGGAGCTCTGAGTGCATGTGTGAGCAACAATGAAGCAGTGTGTGAGTCCATGCTA 2891
DB 2610 CGGAGCTCTGAGTGCATGTGTGAGCAACAATGAAGCAGTGTGTGAGTCCATGCTA 2669
QY 2892 TGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCC 2951
DB 2670 TGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCC 2729
QY 2952 TGAAAATGTTTCAGGCTACTGTACCTGTAGTCAATGTTGGAGCAACAGGCTGTGGCTG 3011
DB 2730 TGAAAATGTTTCAGGCTACTGTACCTGTAGTCAATGTTGGAGCAACAGGCTGTGGCTG 2789
QY 3012 GTGTACTCATCCAGCAATATCTGCAAGAGGAATGCAATAGAGGTTCTTAAAGGACC 3071
DB 2790 GTGTACTCATCCAGCAATATCTGCAAGAGGAATGCAATAGAGGTTCTTAAAGGACC 2849
QY 3072 AGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTTATCCACAGCCCTGCTCAATTC 3131
DB 2850 AGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTTATCCACAGCCCTGCTCAATTC 2909
QY 3132 CAGCATGTCTTAGAGGACAGCATCAACTGGTCTTTCATTTCACTGTCCAGCTTGCCA 3191
DB 2910 CAGCATGTCTTAGAGGACAGCATCAACTGGTCTTTCATTTCACTGTCCAGCTTGCCA 2969
QY 3192 ATGCAACCGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGAC 3251
DB 2970 ATGCAACCGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGAC 3029
QY 3252 CACAGGCAAGCAGTGTGAGACCTGCGATATCTGGCTTCTACGGTATGCCCAATGAGG 3311
DB 3030 CACAGGCAAGCAGTGTGAGACCTGCGATATCTGGCTTCTACGGTATGCCCAATGAGG 3089
QY 3312 GAAATGTGAGGCAATGCAAGTGCATGAGGAGGAGTGTGAGAAAGTGTGAGAACCTGAC 3371
DB 3090 GAAATGTGAGGCAATGCAAGTGCATGAGGAGGAGTGTGAGAAAGTGTGAGAACCTGAC 3149
QY 3372 GTGCTTCTGCAACCAAGGCGCTCAAGGGGAGCAGAGTGCAGCTATGTGAGGTAGAAA 3431
DB 3150 GTGCTTCTGCAACCAAGGCGCTCAAGGGGAGCAGAGTGCAGCTATGTGAGGTAGAAA 3209
QY 3432 TCGATACCAAGGAAACCTCTCAGAGGAAATGTTTATATATCTCTTATTTGACTATCA 3491
DB 3210 TCGATACCAAGGAAACCTCTCAGAGGAAATGTTTATATATCTCTTATTTGACTATCA 3269
QY 3492 GTTCACCTTTTGTCTATCCAGGAAGATGATGCGCTATTATACAGCTATCAATTTTGTGGC 3551

Db	3270	GTTCACCTTTAGTCTATCCAGGAGAGATGATCGCTATTACACAGCTATCAATTTTGTGGC	3329
Qy	3552	TACTCTGAGCAACAAACAGGGATTTGGACATGTTTCATCAATGCCTCCAGAAATTTCAA	3611
Db	3330	TACTCTGAGCAACAAACAGGGATTTGGACATGTTTCATCAATGCCTCCAGAAATTTCAA	3389
Qy	3612	CCTCAACATCAGCTGGGCTGCGCTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCC	3671
Db	3390	CCTCAACATCAGCTGGGCTGCGCTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCC	3449
Qy	3672	TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGA	3731
Db	3450	TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGA	3509
Qy	3732	TTTTGCAACACCAACATATACATTTCTTTTGTATGTAGTAATTTACCTTGGCCCAT	3791
Db	3510	TTTTGCAACACCAACATATACATTTCTTTTGTATGTAGTAATTTACCTTGGCCCAT	3569
Qy	3792	CAAAATTCAGATTCGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTCTTCGT	3851
Db	3570	CAAAATTCAGATTCGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTCTTCGT	3629
Qy	3852	GACTTCTCAGTTGTTTCTCTCTTTGCTCCTGGTGGCTGCTGTGTTTGGAGATCAA	3911
Db	3630	GACTTCTCAGTTGTTTCTCTCTTTGCTCCTGGTGGCTGCTGTGTTTGGAGATCAA	3689
Qy	3912	ACAAAGTTTGGGCTCCAGACGTAGAGACCACTTCTTCGAGAGATGCAACAGATGGC	3971
Db	3690	ACAAAGTTTGGGCTCCAGACGTAGAGACCACTTCTTCGAGAGATGCAACAGATGGC	3749
Qy	3972	CAGCGCTCCCTTTCCTCTGTAATGTGGCTTGGAAACAGATGAGAGCTCTCTGATCT	4031
Db	3750	CAGCGCTCCCTTTCCTCTGTAATGTGGCTTGGAAACAGATGAGAGCTCTCTGATCT	3809
Qy	4032	TATTGGGGGAGTAAAGACTGTTTCCAAACCCATTGCACTGAGCGCTGTTTGGCAA	4091
Db	3810	TATTGGGGGAGTAAAGACTGTTTCCAAACCCATTGCACTGAGCGCTGTTTGGCAA	3869
Qy	4092	CAAAAGCCCTCTCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCC	4151
Db	3870	CAAAAGCCCTCTCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCC	3929
Qy	4152	TCCTGGGAGTCAGGTCCTGTTGGCCAGCGCCCTGGTGACATTTCTAGCAGATGCC	4211
Db	3930	TCCTGGGAGTCAGGTCCTGTTGGCCAGCGCCCTGGTGACATTTCTAGCAGATGCC	3989
Qy	4212	GATAGTGACAAAGAGAGAGTCAGGAGCGCTGAGAAAACCGAAGCAGCAGCCCTCGACA	4271
Db	3990	GATAGTGACAAAGAGAGAGTCAGGAGCGCTGAGAAAACCGAAGCAGCAGCCCTCGACA	4049
Qy	4272	GCCTGGGACCTGCATCTGA	4290
Db	4050	GCCTGGGACCTGCATCTGA	4068
RESULT 5			
AAZ91920			
ID	AAZ91920 standard; cDNA; 8590 BP.		
AC	AAZ91920;		
XX			
DT	08-JUN-2000 (first entry)		
DE	Human mahogany protein coding sequence #2.		
XX			
KW	Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;		
KW	weight regulation; cell therapy; body weight disorder; cachexia;		
KW	anorexia; hyperpigmentation; increased metabolic rate disorder;		
KW	hyperphagia; Antiobesity; antianorexia; anticachexia; ss.		
XX			
OS	Homo sapiens.		
XX			

PN	WO200005373-A2.
XX	
PD	03-FEB-2000.
XX	
PF	21-JUL-1999; 99WO-US016484.
XX	
PR	21-JUL-1998; 98US-0093630P.
PR	20-OCT-1998; 98US-0104978P.
PR	05-FEB-1999; 99US-00245041.
XX	
PA	(MILL-) MILLENIUM PHARM INC.
XX	
PI	Moore K, Nagle DL;
XX	
DR	WPI; 2000-195103/17.
DR	P-PSDB; AAY81807.
XX	
PT	New human and murine mahogany genes, useful, e.g. for diagnosis and treatment of body weight disorders.
PS	Claim 1; Fig 18a; 188pp; English.
CC	This sequence represents a human mahogany gene of the invention. The mahogany genes are used: (i) to produce recombinant mahogany (mg) proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming therapeutics; (iii) as a source of diagnostic probes and primers for detecting expression of mg genes or mutations, regulatory defects, in this gene, or for isolation of related sequences; and (iv) in (cell-based) gene therapy. (ii) are used to raise specific antibodies (Ab); to identify other (extra)cellular products involved in weight regulation, and to screen for agents that disrupt interaction between (ii) and other macromolecules. The Ab are used to detect abnormal levels (or function) of (ii) for diagnosis, prognosis or monitoring of treatment; to evaluate (ii)-expressing cells intended for cell therapy, and as therapeutic mg inhibitors. Cells that express the mg gene (or contain the mg polypeptide) are used to identify agents (A) that modulate mg activity. (A) are potentially useful for the treatment of body weight disorders, particularly obesity, cachexia or anorexia, or other conditions associated with the mg gene such as hyperpigmentation, hyperphagia and disorders that result in increased metabolic rate.
SQ	Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;
Query Match 88.0%; Score 3774; DB 3; Length 8590;	
Best Local Similarity 99.9%; Pred. No. 0;	
Matches 3974; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy	313 TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCCAGTGGCTGTGCCCCCGGCTGG 372
Db	310 TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCCAGTGGCTGTGCCCCCGGCTGG 369
Qy	373 GTGGGCGAGCAATGCCAGCACTGGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 432
Db	370 GTGGGCGAGCAATGCCAGCACTGGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 429
Qy	433 GTGACAGATGGACCTGGAATTAATAACAAACGAAGTCACGCTGGCTCAATTGAAGA 492
Db	430 GTGACAGATGGACCTGGAATTAATAATAACAAACGAAGTCACGCTGGCTCAATTGAAGA 489
Qy	493 CAGCCAAATAGAAATAGAGATTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGAC 552
Db	490 CAGCCAAATAGAAATAGAGATTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGAC 549
Qy	553 CATTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTTCCTGCATTAGTGGC 612
Db	550 CATTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTTCCTGCATTAGTGGC 609
Qy	613 CTCATTGTTTCTTGAGAGATGCAATGAGACTGTCCCTGAGGTGTTTCCACATCAGGT 672
Db	610 CTCATTGTTTCTTGAGAGATGCAATGAGACTGTCCCTGAGGTGTTTCCACATCAGGT 669
Qy	673 TATGCTTCTGCTGCAATTTTTTTTAGTGAAGTCTGTATTAATTTGACTGGATTTATATCT 732

QY 2893 GTGGCCCTCCCTCCCTTTGGCCAGTGATGAATGATATACGATGAGCACCTGCCCCCT 2952
Db |||||
QY 2890 GTGGCCCTCCCTTTGGCCAGTGATGAATGATATACGATGAGCACCTGCCCCCT 2949
Db |||||
QY 2953 GAAATTTGTTTCCAGGCTACTGTACCTGTAGTCAITTCCTTGGAGCAACCCAGGCTGTGGCTGG 3012
Db |||||
QY 2950 GAAATTTGTTTCCAGGCTACTGTACCTGTAGTCAITTCCTTGGAGCAACCCAGGCTGTGGCTGG 3009
Db |||||
QY 3013 TGTACTGATCCCAAGCAATCTGGCAAAAGGAATGATAGAGGTTCCCTAATGAAGACCA 3072
Db |||||
QY 3010 TGTACTGATCCCAAGCAATCTGGCAAAAGGAATGATAGAGGTTCCCTAATGAAGACCA 3069
Db |||||
QY 3073 GTGAAGATGCTTCCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC 3132
Db |||||
QY 3070 GTGAAGATGCTTCCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC 3129
Db |||||
QY 3133 AGCATGTGTAGAGGACAGCAGATACAACTGGTCTTTCTATCTGTCTCAGCTTGGCAA 3192
Db |||||
QY 3130 AGCATGTGTAGAGGACAGCAGATACAACTGGTCTTTCTATCTGTCTCAGCTTGGCAA 3189
Db |||||
QY 3193 TGCACCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 3252
Db |||||
QY 3190 TGCACCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 3249
Db |||||
QY 3253 ACAGCAAGCACTCCGAGACCTGCAATCTGGCTTTCTACGGTATCCCAATGAGGG 3312
Db |||||
QY 3250 ACAGCAAGCACTCCGAGACCTGCAATCTGGCTTTCTACGGTATCCCAATGAGGG 3309
Db |||||
QY 3313 AAATGTACGCATCGAATGAGGCAAGGCTCTGTGCAACACCAACAGGGCAAG 3372
Db |||||
QY 3310 AAATGTACGCATCGAATGAGGCAAGGCTCTGTGCAACACCAACAGGGCAAG 3369
Db |||||
QY 3373 TGCCTTCTGCACCAACCAAGGGCTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAT 3432
Db |||||
QY 3370 TGCCTTCTGCACCAACCAAGGGCTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAT 3429
Db |||||
QY 3433 CGATACAGAAACCCCTCTCAGAGGACATGTTATTAATCTTCTATGACTATCAG 3492
Db |||||
QY 3430 CGATACAGAAACCCCTCTCAGAGGACATGTTATTAATCTTCTATGACTATCAG 3489
Db |||||
QY 3493 TTCACTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATCAATTTGTGGCT 3552
Db |||||
QY 3490 TTCACTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATCAATTTGTGGCT 3549
Db |||||
QY 3553 ACTCTGACGAAACAAACAGGGATTTGACATGTTCTCAATGCTCCCAAGAAATTTCAAC 3612
Db |||||
QY 3550 ACTCTGACGAAACAAACAGGGATTTGACATGTTCTCAATGCTCCCAAGAAATTTCAAC 3609
Db |||||
QY 3613 CTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCT 3672
Db |||||
QY 3610 CTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCT 3669
Db |||||
QY 3673 GTTGTTCCTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGAT 3732
Db |||||
QY 3670 GTTGTTCCTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGAT 3729
Db |||||
QY 3733 TTTCGACACCAACCAATATCACTTTCTTTGTTATGTCAGTAATTTTCACTGGCCATC 3792
Db |||||
QY 3730 TTTCGACACCAACCAATATCACTTTCTTTGTTATGTCAGTAATTTTCACTGGCCATC 3789
Db |||||
QY 3793 AAAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTTCTTGGTG 3852
Db |||||
QY 3790 AAAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTTCTTGGTG 3849
Db |||||
QY 3853 ACTTTCCTTCAGTTGTTTCTCTTCTTGTCTGTGGTGTCTGTGGTGTGAAGATCAAA 3912
Db |||||
QY 3850 ACTTTCCTTCAGTTGTTTCTCTTGTCTGTGGTGTCTGTGGTGTGAAGATCAAA 3909
Db |||||
QY 3913 CAAGTTCTGGGCTCCAGAGTAGAGAGCACTTCTTCAGAGATGCAACAGATGGCC 3972
Db |||||
QY 3910 CAAGTTCTGGGCTCCAGAGTAGAGAGCACTTCTTCAGAGATGCAACAGATGGCC 3969
Db |||||
QY 3973 AGCCGCTCCCTTTGGCTCTGTAATGTGCTGCTTGGAAACAGATGAGGACCTTCCTGATCTT 4032

Db 3970 AGCCGCTCCCTTTGGCTCTGTAATGTGCTTGGAAACAGATGAGGACCTTCCTGATCTT 4029
QY 4033 ATTGGGGGAGTATAAAGACTGTTCCAAAACCCATTGACATGGAGCCGTTGTTGGCAAC 4092
Db 4030 ATTGGGGGAGTATAAAGACTGTTCCAAAACCCATTGACATGGAGCCGTTGTTGGCAAC 4089
QY 4093 AAAGCCGCTGCTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCCT 4152
Db 4090 AAAGCCGCTGCTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCCT 4149
QY 4153 CTTGGGAGTCAAGTCTTGTCTGTGCTGCGCAGCGCCCTGTGTGGACATTTCTCAGCAGATGCGG 4212
Db 4150 CTTGGGAGTCAAGTCTTGTCTGTGCTGCGCAGCGCCCTGTGTGGACATTTCTCAGCAGATGCGG 4209
QY 4213 ATAGTGTACAAGGAGAAGTCAAGAGCCGTGAGAAACCCGAAACAGCAGCCCTCCACAG 4272
Db 4210 ATAGTGTACAAGGAGAAGTCAAGAGCCGTGAGAAACCCGAAACAGCAGCCCTCCACAG 4269
QY 4273 CTTGGGACCTGCATCTGA 4290
Db 4270 CTTGGGACCTGCATCTGA 4287

RESULT 6
AAZ91921
ID AAZ91921 standard; cDNA; 4072 BP.
XX
AC AAZ91921;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #3.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Anticobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
PI WPI; 2000-195103/17.
XX
DR P-PSDB; AAY81808.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 19a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (ii); (iii) as a source of diagnostic probes and primers for
CC therapeutics; (iv) as a source of mutations, regulatory defects, in
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (ii) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)

CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;
Query Match 76.6%; Score 3285; DB 3; Length 4072;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 313 TGTGTCAACGGGGTGGTGGCAACCTGGACCGGCGAGTCGGCTGCGCCCGCGGTGG 372
DB 310 TGTGTCAACGGGGTGGTGGCAACCTGGACCGGCGAGTCGGCTGCGCCCGCGGTGG 369
QY 373 GTGGCGAGCAATGCCAGACTGGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTT 432
DB 370 GTGGCGAGCAATGCCAGACTGGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTT 429
QY 433 GTGACAGATGACCTGGAAATTAATAACAAAGAGTGCAGTGGCTCAATTGAAGGA 492
DB 430 GTGACAGATGACCTGGAAATTAATAACAAAGAGTGCAGTGGCTCAATTGAAGGA 489
QY 493 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGAC 552
DB 490 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGAC 549
QY 553 CATTTATATGTTATGATGGGACTCAATTTATGCAACGCTAGTTGCTGCATTTAGTGGC 612
DB 550 CATTTATATGTTATGATGGGACTCAATTTATGCAACGCTAGTTGCTGCATTTAGTGGC 609
QY 613 CTCATTGTTCTCGAGAGATGGCAATAGAGACTGTCCTGAGGTTGTTGGCACATCAGGT 672
DB 610 CTCATTGTTCTCGAGAGATGGCAATAGAGACTGTCCTGAGGTTGTTGGCACATCAGGT 669
QY 673 TATGCCCTGCTGCATTTTATGATGATGCTGCTTATATTAATTTGACTGGATTAATTAAT 732
DB 670 TATGCCCTGCTGCATTTTATGATGATGCTGCTTATATTAATTTGACTGGATTAATTAAT 729
QY 733 TACAGTTTGTATGATGCTCCAAATAACTGTCAGCGCGAGGAGTGTAAAGTCAAGTAAT 792
DB 730 TACAGTTTGTATGATGCTCCAAATAACTGTCAGCGCGAGGAGTGTAAAGTCAAGTAAT 789
QY 793 AGCAGCGAACTGTTGAAATGTGAAATGTTCTGAAAACTGGAAAGTGAAGCATGTGACATT 852
DB 790 AGCAGCGATCTGTTGAAATGTGAAATGTTCTGAAAACTGGAAAGTGAAGCATGTGACATT 849
QY 853 CCTCACTGTACAGCAACTGTTTTCCTTCATCGAGCATCTGCAATTCAGCTGATGTC 912
DB 850 CCTCACTGTACAGCAACTGTTTTCCTTCATCGAGCATCTGCAATTCAGCTGATGTC 909
QY 913 AGAGGATGCTGCTGCTTCAGACTGGCAGGCTGCTGATGTTTTCAGTTTCTGATACCACT 972
DB 910 AGAGGATGCTGCTGCTTCAGACTGGCAGGCTGCTGATGTTTTCAGTTTCTGATACCACT 969
QY 973 AACAGTCAATTTGGACTCGAGAGGAATATTTCTAACTTAAAGTCTCCCGAGCATCTCAT 1032
DB 970 AACAGTCAATTTGGACTCGAGAGGAATATTTCTAACTTAAAGTCTCCCGAGCATCTCAT 1029
QY 1033 AAAGCTGTGCTCAATGGAACATTTATGTTGGTGTGTTGAGGATATATGTTCAACCACTCA 1092
DB 1030 AAAGCTGTGCTCAATGGAACATTTATGTTGGTGTGTTGAGGATATATGTTCAACCACTCA 1089
QY 1093 GATTATTAACATGTTTCTAGCTATGACCTTCTTATAGGAGTGGCTTCCACTAAACCGT 1152
DB 1090 GATTATTAACATGTTTCTAGCTATGACCTTCTTATAGGAGTGGCTTCCACTAAACCGT 1149
QY 1153 TCTGTGAACAATGTGTTGTTAGATATGCTCATTTTGGCATTTATACAAGGATAAAT 1212

DB 1150 TCTGTGAACAATGTGTTGTTAGATATGCTCATTTCTTTGGCATTATACAAGGATAAAT 1209
QY 1213 TACATGTATGGAGGAAAAATTCATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT 1272
DB 1210 TACATGTATGGAGGAAAAATTCATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT 1269
QY 1273 CACATTCAATAAGTCAATGGGTGTTTGTGACCCCTTAAGGCAAGGAGCAGTATCAGTG 1332
DB 1270 CACATTCAATAAGTCAATGGGTGTTTGTGACCCCTTAAGGCAAGGAGCAGTATCAGTG 1329
QY 1333 GTTGGCACTCTGCACACATTTTACACTGAAGAATGGCCGAGTGGTCATGCTGCTATC 1392
DB 1330 GTTGGCACTCTGCACACATTTTACACTGAAGAATGGCCGAGTGGTCATGCTGCTATC 1389
QY 1393 TTTGGTCACTGCCCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGATAAG 1452
DB 1390 TTTGGTCACTGCCCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGATAAG 1449
QY 1453 AACATGAGATATATACACACCGGTCCTTGTGCAAGGGGTTTACGGCCATAGC 1512
DB 1450 AACATGAGATATATACACACCGGTCCTTGTGCAAGGGGTTTACGGCCATAGC 1509
QY 1513 AGTGTTTAGACCATAGGACAGGCGCTTATACGTTTCATGTTGGCTACAAGGCTTTTCACT 1572
DB 1510 AGTGTTTAGACCATAGGACAGGCGCTTATACGTTTCATGTTGGCTACAAGGCTTTTCACT 1569
QY 1573 GCCAATAAGTACCGGCTTGCAGATGATCTTACCAATATGATGTGGATACCCAGATGTGG 1632
DB 1570 GCCAATAAGTACCGGCTTGCAGATGATCTTACCAATATGATGTGGATACCCAGATGTGG 1629
QY 1633 ACCATTCTTAAGGACAGCGATTTTCCGTTTACCTTGTGACACAGCTGTGATGAGTGA 1692
DB 1630 ACCATTCTTAAGGACAGCGATTTTCCGTTTACCTTGTGACACAGCTGTGATGAGTGA 1689
QY 1693 ACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGGGCGCAAA 1752
DB 1690 ACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGGGCGCAAA 1749
QY 1753 TGCTTCTCTCAGATTTTCAATGCTTACATTTGCTGCTGAGCCGCTGCTGAGTGTGCTCC 1812
DB 1750 TGCTTCTCTCAGATTTTCAATGCTTACATTTGCTGCTGAGCCGCTGCTGAGTGTGCTCC 1809
QY 1813 AGACCTGATCTCCACCATGATGTCAACAGATTTGCGCATTCAGCAGTCTTACCAACAGC 1872
DB 1810 AGACCTGATCTCCACCATGATGTCAACAGATTTGCGCATTCAGCAGTCTTACCAACAGC 1869
QY 1873 ACCATGATGTTTGGTGGTTTCAATAGTCTTCTTCTCAGGACATCTCTGGTATTCACC 1932
DB 1870 ACCATGATGTTTGGTGGTTTCAATAGTCTTCTTCTCAGGACATCTCTGGTATTCACC 1929
QY 1933 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGGCTGTTTGTAGCAGCAGGACCTGCTATT 1992
DB 1930 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGGCTGTTTGTAGCAGCAGGACCTGCTATT 1989
QY 1993 CGGTGTGTGGAACACAGGTCGCTCAGTGTATCTCTGCTGCGGCTGGCAACTGATGAA 2052
DB 1990 CGGTGTGTGGAACACAGGTCGCTCAGTGTATCTCTGCTGCGGCTGGCAACTGATGAA 2049
QY 2053 CAGAAGAAAAGTTAAATCAGATGTTTTCCTTCCAAAGAACTCTTGACCATGACAGATGT 2112
DB 2050 CAGAAGAAAAGTTAAATCAGATGTTTTCCTTCCAAAGAACTCTTGACCATGACAGATGT 2109
QY 2113 GACGAGCACACAGATTTGTTACAGCTGTGACGCAACACCAATGATGCTGCTGCTGCAAT 2172
DB 2110 GACGAGCACACAGATTTGTTACAGCTGTGACGCAACACCAATGATGCTGCTGCTGCAAT 2169
QY 2173 GACCAATGTGCTCCAGGAAACACAGCTGCTCAGNAGCCAGATCTCCATTTTGTAGTAT 2232
DB 2170 GACCAATGTGCTCCAGGAAACACAGCTGCTCAGNAGCCAGATCTCCATTTTGTAGTAT 2229
QY 2233 GAGAATTTGCCCCAAGGATAAAGCCCATGCTACTTGTACTTGTACTGTAAACAAGAGCAGCTGCGAGGAGC 2292
DB 2230 GAGAATTTGCCCCAAGGATAAAGCCCATGCTACTTGTACTTGTACTGTAAACAAGAGCAGCTGCGAGGAGC 2289

XX The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding soluble attractin-1 obtained from
CC activated T cell and human fetal liver libraries. This sequence can be
CC used to enhance immune response in immunosuppressed patients such as
CC those undergoing chemo- and radio-therapy treatment for cancer or those
CC suffering from common variable immunodeficiency syndrome. The proteins
CC may also be used to screen modulators (agonists and antagonists) of
CC immune responses which may also be used to regulate immune reactions.
CC Attractin antibodies can be used to inhibit immune response in transplant
CC recipients or patients afflicted with autoimmune disease
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 71.0%; Score 3045; DB 3; Length 3597;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY	312	CTGTGTCAACGGCGTGCCTCAACCTCGCACCGCCAGTGGCTCTGCCCGCGCGCTG	371
DB	90	CTGTGTCAACGGCGTGCCTCAACCTCGCACCGCCAGTGGCTCTGCCCGCGCGCTG	149
QY	372	GGTGGGAGCAATGCCAGCACTCGGGGGCGGTTTCAAGTAATCGATCTTCTGGGTT	431
DB	150	GGTGGGAGCAATGCCAGCACTCGGGGGCGGTTTCAAGTAATCGATCTTCTGGGTT	209
QY	432	TGTGACAGATGGACTGGAAATTAATACAAACGACGTGCAGTGGCTCATTTGAGG	491
DB	210	TGTGACAGATGGACTGGAAATTAATACAAACGACGTGCAGTGGCTCATTTGAGG	269
QY	492	ACAGCCAAATAGAATATGACACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGA	551
DB	270	ACAGCCAAATAGAATATGACACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGA	329
QY	552	CCATTTATGTTTATGATGGGACTCAATTTATGACCGGTAGTGTCTGCATTTAGTGG	611
DB	330	CCATTTATGTTTATGATGGGACTCAATTTATGACCGGTAGTGTCTGCATTTAGTGG	389
QY	612	CCTCATTTCTCTGAGAGATGCAATGACACTGTCCCTGAGTGTGTGTCACATCAGG	671
DB	390	CCTCATTTCTCTGAGAGATGCAATGACACTGTCCCTGAGTGTGTGTCACATCAGG	449
QY	672	TTATGCCCTGTGCTATTTTATGATGCTGCTTATAATTTGACTGCAATTAATATTAC	731
DB	450	TTATGCCCTGTGCTATTTTATGATGCTGCTTATAATTTGACTGCAATTAATATTAC	509
QY	732	TTACAGTTTTCATATGTCCTCAATTAATCTGTCAGGCGGAGGAGTGAAGTACAGTAA	791
DB	510	TTACAGTTTTCATATGTCCTCAATTAATCTGTCAGGCGGAGGAGTGAAGTACAGTAA	569
QY	792	TAGCAGCGAAACTGTGAATGTGAATGTCTGAAACTGGAAGAGTGAAGCATGTGACAT	851
DB	570	TAGCAGCGAACTGTGTGAATGTGAATGTCTGAAACTGGAAGAGTGAAGCATGTGACAT	629
QY	852	TCCTCACTGTACACAACTGTGTTTTTCTCATCGAGGATCTGCAATTCAGTGTATGT	911
DB	630	TCCTCACTGTACACAACTGTGTTTTTCTCATCGAGGATCTGCAATTCAGTGTATGT	689
QY	912	CAGAGGATGTCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTCACTTCTGACAGC	971
DB	690	CAGAGGATGTCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTCACTTCTGACAGC	749
QY	972	TAAACAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGCATCTCA	1031
DB	750	TAAACAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGCATCTCA	809
QY	1032	TAAAGCTGTGTCAATGGAACATTATGTGGGTTTGGAGGATATATGTTTCAACCACTC	1091
DB	810	TAAAGCTGTGTCAATGGAACATTATGTGGGTTTGGAGGATATATGTTTCAACCACTC	869

QY	1092	AGATTATTAACATGGTCTTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCG	1151
DB	870	AGATTATTAACATGGTCTTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCG	929
QY	1152	TTCTGTGAACAATGTGGTGTGTAGATATGGTCAATCTTTGGCATTATACAGGATAAAT	1211
DB	930	TTCTGTGAACAATGTGGTGTGTAGATATGGTCAATCTTTGGCATTATACAGGATAAAT	989
QY	1212	TTACATGTATGAGAGAAAATTTGATCCAACTGGGAATGTGACCAATGATGTGAGATTTT	1271
DB	990	TTACATGTATGAGAGAAAATTTGATCCAACTGGGAATGTGACCAATGATGTGAGATTTT	1049
QY	1272	TCACATTCATATGATGATCATGGTGTGTGACCCCTAAGGCAAAAGGAGCATGTGACAT	1331
DB	1050	TCACATTCATATGATGATCATGGTGTGTGACCCCTAAGGCAAAAGGAGCATGTGACAT	1109
QY	1332	GGTTGGGCACTCTGCACACATTTTACACTGNAAGATGCGCCAGTGGCTCATGCTGGTCAT	1391
DB	1110	GGTTGGGCACTCTGCACACATTTTACACTGNAAGATGCGCCAGTGGCTCATGCTGGTCAT	1169
QY	1392	CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATCTGCAGGAATATGATTTGGATAA	1451
DB	1170	CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATCTGCAGGAATATGATTTGGATAA	1229
QY	1452	GAACATGTGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTAGGCCCATAG	1511
DB	1230	GAACATGTGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTAGGCCCATAG	1289
QY	1512	CAGTGTTTACGACCACTAGGACCCAGGGCCCTATACCTTTCATGTTGGTCAACAGGCTTT	1571
DB	1290	CAGTGTTTACGACCACTAGGACCCAGGGCCCTATACCTTTCATGTTGGTCAACAGGCTTT	1349
QY	1572	TGCCAATAGTACCCGGTTGACAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG	1631
DB	1350	TGCCAATAGTACCCGGTTGACAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG	1409
QY	1632	GACCACTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGATGTAGTGG	1691
DB	1410	GACCACTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGATGTAGTGG	1469
QY	1692	AACCATGTCTGGTGTGTTGGGGGAAACACACACAATGACACATCTATGAGCCCATGGCCAA	1751
DB	1470	AACCATGTCTGGTGTGTTGGGGGAAACACACACAATGACACATCTATGAGCCCATGGCCAA	1529
QY	1752	ATGCTTCTTTCAGATTTTCATGGCCCTATGACATTTGCCTGTGACCGCTGGTCACTGCTTC	1811
DB	1530	ATGCTTCTTTCAGATTTTCATGGCCCTATGACATTTGCCTGTGACCGCTGGTCACTGCTTC	1589
QY	1812	CAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT - CAGCAGTCTTTACACAACA	1870
DB	1590	CAGACCTGAT - TCCACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTTACACAACA	1648
QY	1871	GCACCATGTATGTGTGGTGTTCATATGTCTCTCTCCTCAGCGACATCTGCTGATTTCA	1930
DB	1649	GCACCATGTATGTGTGGTGTTCATATGTCTCTCTCCTCAGCGACATCTGCTGATTTCA	1708
QY	1931	CCTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTTGTGTTAGCAGCAGCAGTCTGTA	1990
DB	1709	CCTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTTGTGTTAGCAGCAGCAGTCTGTA	1768
QY	1991	TTCCGTGTGTGTGGAACACAGGGTGTCTCAGTGTATCTCGTGGCGGTGGCACTCATGT	2050
DB	1769	TTCCGTGTGTGTGGAACACAGGGTGTCTCAGTGTATCTCGTGGCGGTGGCACTCATGT	1828
QY	2051	TACAGAGAAAGCTTAAATCAGAACTGTTTTTCCAAAGAACTCTTGACCATCACAGAT	2110
DB	1829	TACAGAGAAAGCTTAAATCAGAACTGTTTTTCCAAAGAACTCTTGACCATCACAGAT	1888
QY	2111	GTGACCAGCACACAGATTTTACAGCTGTACAGCTTACAGCCAAACACCAATGCTGCTGCA	2170
DB	1889	GTGACCAGCACACAGATTTTACAGCTGTACAGCTTACAGCCAAACACCAATGCTGCTGCA	1948

QY 2171 ATGACCAATTGTGTCCTCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAGGT 2230
Db 1949 ATGACCAATTGTGTCCTCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAGGT 2008
QY 2231 ATGAGAAATTGCCCCAAGATTAACCCCATGTACTCTTAACAAGAGACACAGCTGCAGGA 2290
Db 2009 ATGAGAAATTGCCCCAAGATTAACCCCATGTACTCTTAACAAGAGACACAGCTGCAGGA 2068
QY 2291 GCTGTGCCCTGGACCAAACTGCCAGTGGAGGCCCGGAATCAGGAATGCAATGCCCTGTC 2350
Db 2069 GCTGTGCCCTGGACCAAACTGCCAGTGGAGGCCCGGAATCAGGAATGCAATGCCCTGTC 2128
QY 2351 CCGAATAATCTGTGGCAATGGCTGGCATTTGGTTGGAACTCATGTTTGAATAATTACTA 2410
Db 2129 CCGAATAATCTGTGGCAATGGCTGGCATTTGGTTGGAACTCATGTTTGAATAATTACTA 2188
QY 2411 CTGCCAAGGAGAAATTATGACAATGCTAAATTTGTTCTGTAGGAACCAAAATGCCCTTTGG 2470
Db 2189 CTGCCAAGGAGAAATTATGACAATGCTAAATTTGTTCTGTAGGAACCAAAATGCCCTTTGG 2248
QY 2471 CTTCCTTTTAAACCCAGAGAGTAGAATTTGCTTAAAGCAGCTGCGAATATGACAGT 2530
Db 2249 CTTCCTTTTAAACCCAGAGAGTAGAATTTGCTTAAAGCAGCTGCGAATATGACAGT 2308
QY 2531 CATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGGTCGGGATCAATG 2590
Db 2309 CATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGGTCGGGATCAATG 2368
QY 2591 TGTCTTACTGTGTGGAGAGATATGTCCTCAATTTTCAAAATAGTTTACTACAGTGGATGC 2650
Db 2369 TGTCTTACTGTGTGGAGAGATATGTCCTCAATTTTCAAAATAGTTTACTACAGTGGATGC 2428
QY 2651 CGTCTGAGCCAGTGTGCTGGAATCTGTGGAATTTTATCAGAACCCAGTACTCGGGGAC 2710
Db 2429 CGTCTGAGCCAGTGTGCTGGAATCTGTGGAATTTTATCAGAACCCAGTACTCGGGGAC 2488
QY 2711 TGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGCCCTGCAAACC 2770
Db 2489 TGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGCCCTGCAAACC 2548
QY 2771 ACAGTGTAGCAGTGTGCGACACCATGTGCTTGGAGCAGCATGTGGAGATGTGACCA 2830
Db 2549 ACAGTGTAGCAGTGTGCGACACCATGTGCTTGGAGCAGCATGTGGAGATGTGACCA 2608
QY 2831 GCGGCAGCTCTGAGTGCATGTGGTGCAACACATGAAGCAGTGTGTGAGTCTCAATGCCT 2890
Db 2609 GCGGCAGCTCTGAGTGCATGTGGTGCAACACATGAAGCAGTGTGTGAGTCTCAATGCCT 2668
QY 2891 ATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCC 2950
Db 2669 ATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCC 2728
QY 2951 CTGAAATTTGTTCAAGGCTACTGTACTGTACTGTGTAATGCTTGGAGCAACAGGCTGTGGCT 3010
Db 2729 CTGAAATTTGTTCAAGGCTACTGTACTGTACTGTGTAATGCTTGGAGCAACAGGCTGTGGCT 2788
QY 3011 GGTGTACTGATCCAGCAATCTGGCAAGGAAATCATAGAGGTTTCTATAAAGGAC 3070
Db 2789 GGTGTACTGATCCAGCAATCTGGCAAGGAAATCATAGAGGTTTCTATAAAGGAC 2848
QY 3071 CAGTGAAGTGCCTTCCGAGCCCTCAGCAAAATTTCTATCCACAGCCCTGCTCAATT 3130
Db 2849 CAGTGAAGTGCCTTCCGAGCCCTCAGCAAAATTTCTATCCACAGCCCTGCTCAATT 2908
QY 3131 CCAGCATGTGTCTAGAGCAGCAGATACACTCGTCTTTTCAATTCAGTGTCCAGCTTGCC 3190
Db 2909 CCAGCATGTGTCTAGAGCAGCAGATACACTCGTCTTTTCAATTCAGTGTCCAGCTTGCC 2968
QY 3191 AATGCAACGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAAATGTGAGAACCTGA 3250
Db 2969 AATGCAACGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAAATGTGAGAACCTGA 3028
QY 3251 CCACAGGCAACGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATGCCACCAATGGAG 3310

Db 3029 CCACAGGCAACGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCCAATGGAG 3089
QY 3311 GGAAATGTGAGCCATGCAAGTGCATATGGGCACGCGCTCTCTGTGCAACCAACACGGGCA 3370
Db 3089 GGAAATGTGAGCCATGCAAGTGCATATGGGCACGCGCTCTCTGTGCAACCAACACGGGCA 3148
QY 3371 AGTGTCTTCTGACCAACCAAGGCGTCAAGGGGAGAGTGCAGCTATGTGAGGTAGAA 3430
Db 3149 AGTGTCTTCTGACCAACCAAGGCGTCAAGGGGAGAGTGCAGCTATGTGAGGTAGAA 3208
QY 3431 ATCGATATCCAAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTTATTTGACTATC 3490
Db 3209 ATCGATATCCAAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTTATTTGACTATC 3268
QY 3491 AGTTCACCTTTAGTCTATCCCAAGAGATGATCGCTATTTACAGCTATCAATTTTGG 3550
Db 3269 AGTTCACCTTTAGTCTATCCCAAGAGATGATCGCTATTTACAGCTATCAATTTTGG 3328
QY 3551 CTACTCTCTGACCAACCAAGGATTTGGACATGTTTCAATGCTTCCCAAGATTTCA 3610
Db 3329 CTACTCTCTGACCAACCAAGGATTTGGACATGTTTCAATGCTTCCCAAGATTTCA 3388
QY 3611 ACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGATGC 3670
Db 3389 ACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGATGC 3448
QY 3671 CTGTTGTTTCAAAACCAACATTAAGGAGTACAAGATAGTTTCTTAATGAGAAGTTG 3730
Db 3449 CTGTTGTTTCAAAACCAACATTAAGGAGTACAAGATAGTTTCTTAATGAGAAGTTG 3508
QY 3731 ATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTGAGTAAATTTCACTGGCCCA 3790
Db 3509 ATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTGAGTAAATTTCACTGGCCCA 3568
QY 3791 TCAAAATTCAG 3801
Db 3569 TCAAAATTCAG 3579

RESULT 8

AAS72657
ID AAS72657 standard; cDNA; 3597 BP.
XX AAS72657;
XX AC
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #8461.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR P-PSDB; ABG08470.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess biodiversity.

2
3
4
5
6
7
8
9
10
11
12

Claim 1: SEO ID NO 8461: 103pp: English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (I) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Very Match 71.0%; Score 3045; DB 5; Length 3597;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy	312	CTGTGTCACGCGCGGTGCTCGAACCCCTGCGACCGGCCAGTGGCTGTGCCCCCGCGGCTG	371
Db	90	CTGTGTCACGCGCGGTGCTCGAACCCCTGCGACCGGCCAGTGGCTGTGCCCCCGCGGCTG	149
Qy	372	GGTGGCGGAGCAATGCCAGCACTGCGGGGGCGGTTTCAGACTAACTGGATCTCTCTGGGTT	431
Db	150	GGTGGCGGAGCAATGCCAGCACTGCGGGGGCGGTTTCAGACTAACTGGATCTCTCTGGGTT	209
Qy	432	TGTGACAGATGGACCTCGGAATTTAATAACAACGGAAGTGCACGTGGCTCAATTGAAGG	491
Db	210	TGTGACAGATGGACCTCGGAATTTAATAACAACGGAAGTGCACGTGGCTCAATTGAAGG	269
Qy	492	ACAGCCCAATAGAAATAATGAGACATCGTTTTCATCTATTTTGTCTACAGAGTGTAGTTGGGA	551
Db	270	ACAGCCCAATAGAAATAATGAGACATCGTTTTCATCTATTTTGTCTACAGAGTGTAGTTGGGA	329
Qy	552	CCATTTATATGTTTATGATGGGGACTCAATTTATGCAACCGCTAGTTGCTGCATTTAGTGG	611
Db	330	CCATTTATATGTTTATGATGGGGACTCAATTTATGCAACCGCTAGTTGCTGCATTTAGTGG	389
Qy	612	CCTCATTTGTTCTGACAGAGATGGCAATGAGACTGTCTCCCTGAGGTTGTTGCCACATCAGG	671
Db	390	CCTCATTTGTTCTGACAGAGATGGCAATGAGACTGTCTCCCTGAGGTTGTTGCCACATCAGG	449
Qy	672	TTATGCTTGTCTGCATTTTTTTTAGTGATGCTGCTTAATAATTCACATGGATTTAAATATTAC	731
Db	450	TTATGCTTGTCTGCATTTTTTTTAGTGATGCTGCTTAATAATTCACATGGATTTAAATATTAC	509
Qy	732	TTACAGTTTTTGATATGCTGCCAAATACTGCTCAGGCCGAGAGAGTGTAAAGATCAGTAA	791
Db	510	TTACAGTTTTTGATATGCTGCCAAATACTGCTCAGGCCGAGAGAGTGTAAAGATCAGTAA	569
Qy	792	TAGCAGCGAAACTGTTGAACTGTGAATCTTCTGAAAACCTGGAAGGTGAAGCATGTGCAT	851
Db	570	TAGCAGCGAATCTGTTGAACTGTGAATCTTCTGAAAACCTGGAAGGTGAAGCATGTGCAT	629
Qy	852	TCTTCACCTGTACAGACAACCTGTGTTTTTCTCATCTGAGGCATCTGCAATTCAGATGATGT	911
Db	630	TCTTCACCTGTACAGACAACCTGTGTTTTTCTCATCTGAGGCATCTGCAATTCAGATGATGT	689

312 CTGTGTCAACGGCGGTGCTGCAACCTGGCACCGGCCAGTGCGTCTGCCCGCCGGCTG 371

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes the need for transparency and accountability in financial reporting.

2. The second part of the document outlines the various methods and techniques used to collect and analyze data. It includes a detailed description of the experimental setup and the procedures followed during the study.

3. The third part of the document presents the results of the study, showing the data collected and the analysis performed. It includes tables and graphs to illustrate the findings.

4. The fourth part of the document discusses the implications of the study and the conclusions drawn from the results. It highlights the significance of the findings and their potential applications.

5. The fifth part of the document provides a summary of the key points discussed throughout the document. It reiterates the main findings and the conclusions drawn from the study.

6. The sixth part of the document includes a list of references to the sources used in the study. It provides a comprehensive overview of the literature related to the topic.

7. The seventh part of the document contains a list of figures and tables, providing a visual representation of the data and results.

8. The eighth part of the document includes a list of appendices, providing additional information and data that support the main findings of the study.

9. The ninth part of the document contains a list of footnotes, providing further details and clarifications on specific points discussed in the text.

10. The tenth part of the document includes a list of acknowledgments, thanking the individuals and organizations that provided support and assistance during the study.

11. The eleventh part of the document contains a list of references, providing a comprehensive overview of the literature related to the topic.

12. The twelfth part of the document includes a list of figures and tables, providing a visual representation of the data and results.

13. The thirteenth part of the document contains a list of appendices, providing additional information and data that support the main findings of the study.

14. The fourteenth part of the document includes a list of footnotes, providing further details and clarifications on specific points discussed in the text.

15. The fifteenth part of the document contains a list of acknowledgments, thanking the individuals and organizations that provided support and assistance during the study.

16. The sixteenth part of the document includes a list of references, providing a comprehensive overview of the literature related to the topic.

17. The seventeenth part of the document contains a list of figures and tables, providing a visual representation of the data and results.

18. The eighteenth part of the document includes a list of appendices, providing additional information and data that support the main findings of the study.

19. The nineteenth part of the document contains a list of footnotes, providing further details and clarifications on specific points discussed in the text.

20. The twentieth part of the document includes a list of acknowledgments, thanking the individuals and organizations that provided support and assistance during the study.

D_b 90 CTGTGTCAACGGCGGTGGTCTGCAACCTGGCACCGGCCAGTGGCTCTGCCCGCGGCTG 149

0x 373 CCGGGGCGAATGCCGCACTCCGCCTTCACGGTT 431

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

Db 150 GTGGCGAGCAATGCCAGCACTCGGGGCGGCTTCAGACTAACTGGATCTTCTGGGT 209

[illegible]

432 TGTGACAGATGGACCTGGAAATATATAAATACAAAACGAAGTGCACGTGCTCATTTGAAGG 491
QY

db 210 TGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAAGTGCACGTGGCTCATTTGAAGG 269

QY 492 ACAGCCAAATAGAAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGA 551

270 ACCGCCAATAATCATTCCTTCAATCTTTCTACACTGTAATTCCA 329

[illegible]

QY 552 CCATTATATGTTTATGATGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGG 611

[illegible]

D6 330 CCATTATATGTTTATGATGGGACTCAATTATATGCACCGCTAGTTGCTGCATTTAGTGG 389

612 CCTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGG 671

D_b 390 CCTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGG 449

1769 TTGGTGTGTGTGGAAACACAGGGTCGTCTCAGTGTATCTCTGTGGCGCTGGCACTGATG 1828
2051 AACAGAAAGAAAGTTTAAATCAGATGTTTTCCTCAAAAGAACTCTTGACCATGACAGAT 2110
1829 AACAGAAAGAAAGTTTAAATCAGATGTTTTCCTCAAAAGAACTCTTGACCATGACAGAT 1888
2111 GTGACCAACACACAGATGTTTACAGCTGTGACAGCAACACCAATGACTGCCACCTGGTGCA 2170
1889 GTGACCAACACACAGATGTTTACAGCTGTGACAGCAACACCAATGACTGCCACCTGGTGCA 1948
2171 ATGACCAATGTTGTCCTCCAGGAACACACAGCTGCTCAGAAAGCCAGATCTCCATTTTATAGT 2230
1949 ATGACCAATGTTGTCCTCCAGGAACACACAGCTGCTCAGAAAGCCAGATCTCCATTTTATAGT 2008
2231 ATGAGAAATGTCCTCCAGGAACACCACTGACTACTGTAAACAGACACAGCTGACAGGA 2290
2009 ATGAGAAATGTCCTCCAGGAACACCACTGACTACTGTAAACAGACACAGCTGACAGGA 2068
2291 GCTGTGCCCTGGACCAAGCACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATTGCCCTGC 2350
2069 GCTGTGCCCTGGACCAAGCACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATTGCCCTGC 2128
2351 CCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTGAAATTAATA 2410
2129 CCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTGAAATTAATA 2188
2411 CTGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAACAATGCCCTTTGG 2470
2189 CTGCCAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAACAATGCCCTTTGG 2248
2471 CTTCTCTTAAACCCAGAAAGTAGAATTTGCTTTAAGCAGCTGCGGAATATAGCAGT 2530
2249 CTTCTCTTAAACCCAGAAAGTAGAATTTGCTTTAAGCAGCTGCGGAATATAGCAGT 2308
2531 CATCTCAGAGATGTCAGAGCTCACTTAAACCCATGGCTGGCTTTCGGAAGATCAATG 2590
2309 CATCTCAGAGATGTCAGAGCTCACTTAAACCCATGGCTGGCTTTCGGAAGATCAATG 2368
2591 TGTCTACTGTGTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATGC 2650
2369 TGTCTACTGTGTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATGC 2428
2651 CGTCTGACCCAGATGCTGATCTGTGGAATTTATCAGAACCCAGTACTCGGGAC 2710
2429 CGTCTGACCCAGATGCTGATCTGTGGAATTTATCAGAACCCAGTACTCGGGAC 2488
2711 TGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACC 2770
2489 TGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACC 2548
2771 ACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGACAGATGTGGAGATTGCACCA 2830
2549 ACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGACAGATGTGGAGATTGCACCA 2608
2831 GCGGCAGCTCTGAGTGCATGTGGTGACAGACATGAGCAGTGTGACCTCCCAATGCCT 2890
2609 GCGGCAGCTCTGAGTGCATGTGGTGACAGACATGAGCAGTGTGACCTCCCAATGCCT 2668
2891 ATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCC 2950
2669 ATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCC 2728
2951 CTGAAATGTTTCAGGCTACTGTACTCTGATCATTTGCTTGGAGCAACCGGCTGTGGCT 3010
2729 CTGAAATGTTTCAGGCTACTGTACTCTGATCATTTGCTTGGAGCAACCGGCTGTGGCT 2788
3011 GGTGTACTGATCCAGCAATATGCGCAAGGGAATGATAGAGGGTTCTCTATTAAGGAC 3070
2789 GGTGTACTGATCCAGCAATATGCGCAAGGGAATGATAGAGGGTTCTCTATTAAGGAC 2848
3071 CAGTGAAGATGCTTCGCAAGCCCTCAAGGAAATTTCTATCCACAGCCCTGCTCAATT 3130

2849 CAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTTATCCACAGCCCTGCTCAATT 2908
3131 CCAGCATGTGTCTAGAGGACACAGATCAAACTGGTCTTTTCAATTCAGTCTCCAGTTGGC 3190
2909 CCAGCATGTGTCTAGAGGACACAGATCAAACTGGTCTTTTCAATTCAGTCTCCAGTTGGC 2968
3191 AATGCAACGGCCACAGTAATAATCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGA 3250
2969 AATGCAACGGCCACAGTAATAATCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGA 3028
3251 CCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACCGTGTATCCCAATGGAG 3310
3029 CCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACCGTGTATCCCAATGGAG 3088
3311 GGAATGTCTCAGGCATGCAAGTGGCAAGTGGCAAGTGGCTCTCTGTGCAACCAACACGGGCA 3370
3089 GGAATGTCTCAGGCATGCAAGTGGCAAGTGGCAAGTGGCTCTCTGTGCAACCAACACGGGCA 3148
3371 AGTGTCTTCTGACCAACCAAGGGGCTCAAGGGGGAAGAGTCCAGCTATGTGAGGTAGAA 3430
3149 AGTGTCTTCTGACCAACCAAGGGGCTCAAGGGGGAAGAGTCCAGCTATGTGAGGTAGAA 3208
3431 ATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTGACTATC 3490
3209 ATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTGACTATC 3268
3491 AGTTCACCTTTAGTCTATCCAGGAAGATGATGCTTATTATACAGCTATCAATTTTGTGG 3550
3269 AGTTCACCTTTAGTCTATCCAGGAAGATGATGCTTATTATACAGCTATCAATTTTGTGG 3328
3551 CTACTCTCAGCAACCAACCAAGGATTTGGACATGTTTCATCAATGCTCCAAAGATTTCA 3610
3329 CTACTCTCAGCAACCAACCAAGGATTTGGACATGTTTCATCAATGCTCCAAAGATTTCA 3388
3611 ACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGC 3670
3389 ACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGC 3448
3671 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAGATAGTTTCTCTAATGAGAAGTTG 3730
3449 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAGATAGTTTCTCTAATGAGAAGTTG 3508
3731 ATTTTCGCAACCAACCAATATCAGCTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCA 3790
3509 ATTTTCGCAACCAACCAATATCAGCTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCA 3568
3791 TCAAAATTCAG 3801
3569 TCAAAATTCAG 3579

RESULT, 9
ADD70997
ID ADD70997 standard; DNA; 3597 BP.
XX
AC ADD70997;
XX
DT 15-JAN-2004 (first entry)
XX
Human attractin gene SEQ ID NO:1.
DE
XX
XX liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
XX cytosstatic; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
FN WO2003061564-A2.
PD 31-JUL-2003.
XX
XX 20-DEC-2002; 2002WO-US040718.
PF
XX 21-DEC-2001; 2001US-0341815P.
PR

QY 1812 CAGACCTGATCTCCACCAGATGTAACAGATTGCGCATT-CAGCAGTCTTACAAACA 1870
DB 1590 CAGACCTGAT-TCCACCAGATGTAACAGATTGCGCATTCCAGCAGTCTTACAAACA 1648
QY 1871 GCACCATGATGTTGCGTGGTTCATAGTCTCCCTCAGCGACATCTGGTATTCA 1930
DB 1649 GCACCATGATGTTGCGTGGTTCATAGTCTCCCTCAGCGACATCTGGTATTCA 1708
QY 1931 CCTCGGAACAGTGTGATCGCATCGAGTGAAGCGCTGTTTATAGCAGGACCTGGTA 1990
DB 1709 CCTCGGAACAGTGTGATCGCATCGAGTGAAGCGCTGTTTATAGCAGGACCTGGTA 1768
QY 1991 TTCGGTGTGTGGNACACAGGGTGGTCTCAGTGTATCTCTGGCGCGCTGCACTGATG 2050
DB 1769 TTCGGTGTGTGGNACACAGGGTGGTCTCAGTGTATCTCTGGCGCGCTGCACTGATG 1828
QY 2051 AACAGAAAGAAAAGTTAAATTCAGAAATCTTTTCCAAAAGAACTCTTGACCATGACAGAT 2110
DB 1829 AACAGAAAGAAAAGTTAAATTCAGAAATCTTTTCCAAAAGAACTCTTGACCATGACAGAT 1888
QY 2111 GTGACACGACACAGATTGTTACAGTGTACAGCCNACACCAATGATGCGCACTGGTGCA 2170
DB 1889 GTGACACGACACAGATTGTTACAGTGTACAGCCNACACCAATGATGCGCACTGGTGCA 1948
QY 2171 ATGACCAATTGTCTCCAGGAACACAGCTGCTCAGAAAGGCGAGATCTCCATTTTAGGT 2230
DB 1949 ATGACCAATTGTCTCCAGGAACACAGCTGCTCAGAAAGGCGAGATCTCCATTTTAGGT 2008
QY 2231 ATGAGAAATTGCCAAAGGATAACCCATGTACTACTGTAAACAGAACCCAGCTGCAGGA 2290
DB 2009 ATGAGAAATTGCCAAAGGATAACCCATGTACTACTGTAAACAGAACCCAGCTGCAGGA 2068
QY 2291 GCTGTGCCCTGGACACAGAACTGCCAGTGGAGGCCCGGAAATCAGAGTGCATGCCCTGC 2350
DB 2069 GCTGTGCCCTGGACACAGAACTGCCAGTGGAGGCCCGGAAATCAGAGTGCATGCCCTGC 2128
QY 2351 CGGAAAATATCTGTGGCAATTGGCTGGCAATTTGGTTGAAACTCATGTTTGAATTTACTA 2410
DB 2129 CGGAAAATATCTGTGGCAATTGGCTGGCAATTTGGTTGAAACTCATGTTTGAATTTACTA 2188
QY 2411 CTGCCAAGGAAATATGACAAATGCTAAATGTTCTGTAGGACACCAATGCCCTTTGG 2470
DB 2189 CTGCCAAGGAAATATGACAAATGCTAAATGTTCTGTAGGACACCAATGCCCTTTGG 2248
QY 2471 CTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAAATATGCAAT 2530
DB 2249 CTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAAATATGCAAT 2308
QY 2531 CATCTCAGAGCATGTCCAGCTCACTTAAACCCATGGGTGGCTTCGGAAGATCAATG 2590
DB 2309 CATCTCAGAGCATGTCCAGCTCACTTAAACCCATGGGTGGCTTCGGAAGATCAATG 2368
QY 2591 TGTCTACTGGTCTGGGAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATGC 2650
DB 2369 TGTCTACTGGTCTGGGAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATGC 2428
QY 2651 CGTCTGAGCCCAAGTATGCTGGATCTGTGGAAATTTTATCAGAACCCAGTACTCGGGAC 2710
DB 2429 CGTCTGAGCCCAAGTATGCTGGATCTGTGGAAATTTTATCAGAACCCAGTACTCGGGAC 2488
QY 2711 TGAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCGCTGCAAAACC 2770
DB 2489 TGAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCGCTGCAAAACC 2548
QY 2771 ACAGTCTGAAGCAGTCCCGACACCATGTGCTTGAAGCAGCAGTGTGAGATTGCACCA 2830
DB 2549 ACAGTCTGAAGCAGTCCCGACACCATGTGCTTGAAGCAGCAGTGTGAGATTGCACCA 2608
QY 2831 GCGGCAGCTCTGAGTGCATGTGGTGAGCAAAATGAAAGCAGTGTGGTCAATGCTT 2890
DB 2609 GCGGCAGCTCTGAGTGCATGTGGTGAGCAAAATGAAAGCAGTGTGGTCAATGCTT 2668
QY 2891 ATGTGCGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCC 2950

DB 2669 ATGTGCGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGAGCACTGCCCCC 2728
QY 2951 CTGAAATTTGTTTCAGGCTACTGTACTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGT 3010
DB 2729 CTGAAATTTGTTTCAGGCTACTGTACTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGT 2788
QY 3011 GGTGTACTGATCCAGCAATATCTGCAAAAGGGAATGATAGAGGTTCTTATAAGGAC 3070
DB 2789 GGTGTACTGATCCAGCAATATCTGCAAAAGGGAATGATAGAGGTTCTTATAAGGAC 2848
QY 3071 CAGTGAAGATGCTTCCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 3130
DB 2849 CAGTGAAGATGCTTCCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 2908
QY 3131 CAGCATGTGTCTAGAGGACAGAGATCAACTGCTCTTTCATTCTCTGTCAGCTTGGC 3190
DB 2909 CCAGCATGTGTCTAGAGGACAGAGATCAACTGCTCTTTCATTCTCTGTCAGCTTGGC 2968
QY 3191 AATGCAAGGCCACAGCTAAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGA 3250
DB 2969 AATGCAAGGCCACAGCTAAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGA 3028
QY 3251 CCACAGCAAGCACTCGGAGACCTGATATCTGGCTTCTACGGTGTATCCCAATGGAG 3310
DB 3029 CCACAGCAAGCACTCGGAGACCTGATATCTGGCTTCTACGGTGTATCCCAATGGAG 3088
QY 3311 GGAATGCTAGCCATCAAGTGCATGCGCAATGGGACGCTCTCTGTGCAACACCAACCGGCA 3370
DB 3089 GGAATGCTAGCCATCAAGTGCATGCGCAATGGGACGCTCTCTGTGCAACACCAACCGGCA 3148
QY 3371 AGTGCCTTCTGACACCAACGAGGCGTCAAGGGGACGAGTGCAGCTATGTGAGGTGAAA 3430
DB 3149 AGTGCCTTCTGACACCAACGAGGCGTCAAGGGGACGAGTGCAGCTATGTGAGGTGAAA 3208
QY 3431 ATGATACCAAGAAACCCCTCTCAGAGGACATGTTATTATCTCTTATTGACTATC 3490
DB 3209 ATGATACCAAGAAACCCCTCTCAGAGGACATGTTATTATCTCTTATTGACTATC 3268
QY 3491 AGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGAGTATCAATTTTGTG 3550
DB 3269 AGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGAGTATCAATTTTGTG 3328
QY 3551 CTACTCTGACGAAACAAACAGAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTCA 3610
DB 3329 CTACTCTGACGAAACAAACAGAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTCA 3388
QY 3611 ACCTCAACATCAGCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGC 3670
DB 3389 ACCTCAACATCAGCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGC 3448
QY 3671 CTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTG 3730
DB 3449 CTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTG 3508
QY 3731 ATTTTCGAACCCCAATATCACITTTCTTTTATGTCAGTAAATTTCCACCTGGCCCA 3790
DB 3509 ATTTTCGAACCCCAATATCACITTTCTTTTATGTCAGTAAATTTCCACCTGGCCCA 3568
QY 3791 TCAAAATTCAG 3801
DB 3569 TCAAAATTCAG 3579

RESULT 10

AAZ91922
ID AAZ91922 standard; cDNA; 2625 BP.

XX AAZ91922;

AC AAZ91922;

XX 08-JUN-2000 (first entry)

XX Human mahogany protein coding sequence #4.

DE

XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX Homo sapiens.
XX WO200005373-A2.
XX 03-FEB-2000.
XX 21-JUL-1999; 99WO-US016484.
XX 21-JUL-1998; 98US-0093630P.
XX 20-OCT-1998; 98US-0104978P.
XX 05-FEB-1999; 99US-00245041.
XX (MILL-) MILLINIUM PHARM INC.
XX Moore K, Nagle DL;
XX WPI; 2000-195103/17.
XX P-PSDB; AAY81809.
XX New human and murine mahogany genes, useful, e.g. for diagnosis and
XX treatment of body weight disorders.
XX Claim 1; Fig 20a; 188pp; English.
XX This sequence represents a human mahogany gene of the invention. The
XX mahogany genes are used: (i) to produce recombinant mahogany (mg)
XX proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming
XX therapeutics; (iii) as a source of diagnostic probes and primers for
XX detecting expression of mg genes or mutations, regulatory defects, in
XX this gene, or for isolation of related sequences; and (iv) in (cell-
XX based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
XX identify other (extra)cellular products involved in weight regulation,
XX and to screen for agents that disrupt interaction between (ii) and other
XX macromolecules. The Ab are used to detect abnormal levels (or function)
XX of (ii) (for diagnosis, prognosis or monitoring of treatment); to
XX evaluate (ii)-expressing cells intended for cell therapy, and as
XX therapeutic mg inhibitors. Cells that express the mg gene (or contain the
XX mg polypeptide) are used to identify agents (A) that modulate mg
XX activity. (A) are potentially useful for the treatment of body weight
XX disorders, particularly obesity, cachexia or anorexia, or other
XX conditions associated with the mg gene such as hyperpigmentation,
XX hyperphagia and disorders that result in increased metabolic rate
XX
XX Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;
Query Match 42.8%; Score 1837; DB 3; Length 2625;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2037; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 313 TGTGTCAACGGCGGTGCTGCAACCTGGCACCGGCGAGTGGCTGCCCGCGGCTGG 372
DB 310 TGTGTCAACGGCGGTGCTGCAACCTGGCACCGGCGAGTGGCTGCCCGCGGCTGG 369
QY 373 GTGGCGGAGCAATGCCAGCACTGCCGGGGCGGCTTCAGACTAACTGGATCTTCGGGTTT 432
DB 370 GTGGCGGAGCAATGCCAGCACTGCCGGGGCGGCTTCAGACTAACTGGATCTTCGGGTTT 429
QY 433 GTGACAGATGGACCTGGAAATTTAAATACAAAGAGTGCAGTGGCTCAATTGAAGGA 492
DB 430 GTGACAGATGGACCTGGAAATTTAAATACAAAGAGTGCAGTGGCTCAATTGAAGGA 489
QY 493 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGAC 552
DB 490 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGAC 549
QY 553 CATTATATGTTTATGATGGGAGCTCAATTTATGCAACCGCTAGTGTGCAATTTAGTGGC 612

DB 550 CATTATATGTTTATGATGGGAGCTCAATTTATGCAACCGCTAGTGTGCAATTTAGTGGC 609
QY 613 CTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 672
DB 610 CTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 669
QY 673 TATGCCCTTGCTGCATTTTCTGATGATGCTGCTTATATTTGACTGCTTAAATATTAAT 732
DB 670 TATGCCCTTGCTGCATTTTCTGATGATGCTGCTTATATTTGATTTGATTTAATATTAAT 729
QY 733 TACAGTTTTCATATGTTCTCCAAATAACTGCTCAGGCGGAGGAGAGTGAAGATCAGTAAT 792
DB 730 TACAGTTTTCATATGTTCTCCAAATAACTGCTCAGGCGGAGGAGAGTGAAGATCAGTAAT 789
QY 793 AGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGATGTGACATT 852
DB 790 AGCAGCGATATCTGTTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGATGTGACATT 849
QY 853 CCTCAGCTGTACAGCAACTGTGTTTCTCTCATCGAGGCACTGTGCAATTCAAGTGATGTC 912
DB 850 CCTCAGCTGTACAGCAACTGTGTTTCTCTCATCGAGGCACTGTGCAATTCAAGTGATGTC 909
QY 913 AGAGGATGCTCCTGCTCTCAGACTGGCAGGCTCTGGATGTTTCAGTTCCTGTACAGT 972
DB 910 AGAGGATGCTCCTGCTCTCAGACTGGCAGGCTCTGGATGTTTCAGTTCCTGTACAGT 969
QY 973 AACCACTGTTTGGACTCGAGAGGAATATCTAACTTAAAGCTCCCGAGAGCACTCAT 1032
DB 970 AACCACTGTTTGGACTCGAGAGGAATATCTAACTTAAAGCTCCCGAGAGCACTCAT 1029
QY 1033 AAAGCTGTGCTCAATCGAAACATTTATGTTGGTGTGTTGGAGGATATATGTTCAACCACTCA 1092
DB 1030 AAAGCTGTGCTCAATCGAAACATTTATGTTGGTGTGTTGGAGGATATATGTTCAACCACTCA 1089
QY 1093 GATTATAACATGTTCTAGGATGACCTTCTTAGGGAGTGGCTTCACATAACCT 1152
DB 1090 GATTATAACATGTTCTAGGATGACCTTCTTAGGGAGTGGCTTCACATAACCT 1149
QY 1153 TCTGTGCAACAATGTTGTTTGTAGATGTTGCTTCTTGGCATTTACAAAGATATAAAT 1212
DB 1150 TCTGTGCAACAATGTTGTTTGTAGATGTTGCTTCTTGGCATTTACAAAGATATAAAT 1209
QY 1213 TACATGTTGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGATTTT 1272
DB 1210 TACATGTTGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGATTTT 1269
QY 1273 CACATTCATAATGAGTCATGGGTGTTTGAACCCCTTAAGGCAAGAGCAGTATGACGTG 1332
DB 1270 CACATTCATAATGAGTCATGGGTGTTTGAACCCCTTAAGGCAAGAGCAGTATGACGTG 1329
QY 1333 GTTGGCACTCTGCAACACATTTTACACTGAAGATGGCCGAGTGGTCAATGCTGTGATC 1392
DB 1330 GTTGGCACTCTGCAACACATTTTACACTGAAGATGGCCGAGTGGTCAATGCTGTGATC 1389
QY 1393 TTTGTTCACTGCCCTCTCTATGATATATAAGCAATGTGACGAATATGATTTGGATAG 1452
DB 1390 TTTGTTCACTGCCCTCTCTATGATATATAAGCAATGTGACGAATATGATTTGGATAG 1449
QY 1453 AACACATGGAGTATATTACACACCCAGGCTGCCCTTGTGCAAGGGGTTTACGCCCATAGC 1512
DB 1450 AACACATGGAGTATATTACACACCCAGGCTGCCCTTGTGCAAGGGGTTTACGCCCATAGC 1509
QY 1513 AGTGTTTACGACATAGGACACAGGCGCTTATACGTTTCATGTTGGCTACAAAGCTTTTCA 1572
DB 1510 AGTGTTTACGACATAGGACACAGGCGCTTATACGTTTCATGTTGGCTACAAAGCTTTTCA 1569
QY 1573 GCCAATAAGTACCGGCTTGCAGATGATCTTACCATATGATGTTGGATACCCAGATGTTGG 1632
DB 1570 GCCAATAAGTACCGGCTTGCAGATGATCTTACCATATGATGTTGGATACCCAGATGTTGG 1629
QY 1633 ACCATTTTAAAGGACACAGCGATTTTTCGTTTCTGTTGCAACAGCTGTGATGATGATGGA 1692
DB 1630 ACCATTTTAAAGGACACAGCGATTTTTCGTTTCTGTTGCAACAGCTGTGATGATGATGGA 1689

DB 983 CATCAATCAGACATCTGTGAGAGTGTGAGAACTGACCAAGCAGCACTGGAGAC 1042
QY 3273 CTGCATATCTGGCTTCTACGGTATCCCAACAAATGGAGGAAATGTACGCCATCAAGTG 3332
DB 1043 CTGCATATCTGGCTTCTACGGTATCCCAACAAATGGAGGAAATGTACGCCATCAAGTG 1102
QY 3333 CAATGGGACGCGTCTCTGTGCAACCAACCAACGCGGAGTCTTCGACCAACCAAGG 3392
DB 1103 CAATGGGACGCGTCTCTGTGCAACCAACCAACGCGGAGTCTTCGACCAACCAAGG 1162
QY 3393 COTCAAGGGGACGAGTCCAGCTATGTGAGTAGAAAATCGATACCAAGAAACCCCTCT 3452
DB 1163 COTCAAGGGGACGAGTCCAGCTATGTGAGTAGAAAATCGATACCAAGAAACCCCTCT 1222
QY 3453 CAGAGAAATGTTATATATATCTCTTATTGACTATCAGTTACGTTTAGTCTATCCCA 3512
DB 1223 CAGAGAAATGTTATATATATCTCTTATTGACTATCAGTTACGTTTAGTCTATCCCA 1282
QY 3513 GGAAGATGATCGCTATACAGAGTATCAATTTTGTGGTACTCTCGACGAAACAAACAG 3572
DB 1283 GGAAGATGATCGCTATACAGAGTATCAATTTTGTGGTACTCTCGACGAAACAAACAG 1342
QY 3573 GGAATTGGACATGTTATCAATGCTTCAAGAAATTTCAACCTCAACATCAGCTGGGTGC 3632
DB 1343 GGAATTGGACATGTTATCAATGCTTCAAGAAATTTCAACCTCAACATCAGCTGGGTGC 1402
QY 3633 CAGTTTCTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTCTTTCAAAACCAACAT 3692
DB 1403 CAGTTTCTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTCTTTCAAAACCAACAT 1462
QY 3693 TAAGGAGTACAAAGATGTTCTTAATGAGAAGTTGATTTTCGAAACCAACCAATAT 3752
DB 1463 TAAGGAGTACAAAGATGTTCTTAATGAGAAGTTGATTTTCGAAACCAACCAATAT 1522
QY 3753 CACTTCTCTGTTATGTCAGTAATTTCACTGCGCCATCAAAATTCAGATGCTCTC 3812
DB 1523 CACTTCTCTGTTATGTCAGTAATTTCACTGCGCCATCAAAATTCAGATGCTCTC 1582
QY 3813 TCAGCACAGCAATTTATGACCTGTCAGTCTTCTGCTGACTTCTTCTAGTTCTTCT 3872
DB 1583 TCAGCACAGCAATTTATGACCTGTCAGTCTTCTGCTGACTTCTTCTAGTTCTTCT 1642
QY 3873 CTCCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3932
DB 1643 CTCCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702
QY 3933 ACCTAGAGAGCACTTCTTCAGAGATGCAACAGATGCGCCGCTGCTGCTGCTGCTGCT 3992
DB 1703 ACCTAGAGAGCACTTCTTCAGAGATGCAACAGATGCGCCGCTGCTGCTGCTGCTGCT 1762
QY 3993 AAATGTGCTGCTGGAACAGATGAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4052
DB 1763 AAATGTGCTGCTGGAACAGATGAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1822
QY 4053 TGTTCCTCAACCCATTCAGTGGAGCGGTGTTTGGCAACAAAGCCGCTGCTGCTGCTGCT 4112
DB 1823 TGTTCCTCAACCCATTCAGTGGAGCGGTGTTTGGCAACAAAGCCGCTGCTGCTGCTGCT 1882
QY 4113 GTTGTGAGGCTCCTCGAGCCCTGGGTTGGATCCCTCTCTCTGGGAGTCAGGCTCTGTC 4172
DB 1883 GTTGTGAGGCTCCTCGAGCCCTGGGTTGGATCCCTCTCTCTGGGAGTCAGGCTCTGTC 1942
QY 4173 TGTGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4232
DB 1943 TGTGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2002
QY 4233 AGGAGCGCTGAGAAACCGGAGCAGCAGCCCTGTCAGCAGCCTGGGACCTGCTGCTGCT 4290
DB 2003 AGGAGCGCTGAGAAACCGGAGCAGCAGCCCTGTCAGCAGCCTGGGACCTGCTGCTGCT 2060

AAS72660
ID AAS72660 standard; cDNA; 3490 BP.
XX
AC AAS72660;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8464.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
XX 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG08473.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 8464; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (I) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Query Match 25.1%; Score 1077; DB 5; Length 3490;

Best Local Similarity 99.6%; Pred No. 0;

Matches 1587; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 409 AGACTAACTGGATCTCTCGGTTTGTGACATGGACCTGGAAATTAATAACAACG 468

DB 22 AGACTAACTGGATCTCTCGGTTTGTGACATGGACCTGGAAATTAATAACAACG 81

QY 469 AAGTCACGTGCTCATTTGAAGGACAGCAATAGAATAATGAGACTTCGTTTCAATCAT 528

DB 82 AAGTCACGTGCTCATTTGAAGGACAGCAATAGAATAATGAGACTTCGTTTCAATCAT 141

CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3165 BP; 1042 A; 705 C; 687 G; 731 T; 0 U; 0 Other;

Query Match 10.0%; Score 428; DB 5; Length 3165;
 Best Local Similarity 100.0%; Pred. No. 2.7e-189;
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2757 AAGGCTGCAACACAGTGTAGCAGTGGCGGACACCATGTGCTTGAGCAGCATG 2816
 DB |||||
 QY 1218 AAGGCTGCAACACAGTGTAGCAGTGGCGGACACCATGTGCTTGAGCAGCATG 1277
 DB |||||
 QY 2817 TGGAGATTGCACCGCGCAGCTGTAGTGTGGTGCAGCAACATGAAGCAGTGTGT 2876
 DB |||||
 QY 1278 TGGAGATTGCACCGCGCAGCTGTAGTGTGGTGCAGCAACATGAAGCAGTGTGT 1337
 DB |||||
 QY 2877 GGACTCCCAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGTTACGAT 2936
 DB |||||
 QY 1338 GGACTCCCAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGTTACGAT 1397
 DB |||||
 QY 2937 GAGCAGCTGCCCCCTGAAATTTTCAGGCTACTGTACCTGTAGTCTATGCTTGGAGCA 2996
 DB |||||
 QY 1398 GAGCAGCTGCCCCCTGAAATTTTCAGGCTACTGTACCTGTAGTCTATGCTTGGAGCA 1457
 DB |||||
 QY 2997 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGGG 3056
 DB |||||
 QY 1458 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGGG 1517
 DB |||||
 QY 3057 TTCCTATAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 3116
 DB |||||
 QY 1518 TTCCTATAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 1577
 DB |||||
 QY 3117 GTCGTCAG 3184
 DB |||||
 QY 1638 CTGTCAG 1645

RESULT 14

AAS76294
 ID AAS76294 standard; cDNA; 3313 BP.

XX
 AC AAS76294;

XX
 DT 13-FEB-2002 (first entry)

XX
 DE DNA encoding novel human diagnostic protein #12098.

XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; sa.

XX
 OS Homo sapiens.

XX
 FN WO200175067-A2.

XX
 PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 XX
 PI WPI; 2001-639362/73.
 DR P-PSDB; ABG12107.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 XX Claim 1; SEQ ID NO 12098; 103pp: English.
 PS
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3313 BP; 1050 A; 731 C; 724 G; 768 T; 0 U; 0 Other;
 Query Match 10.0%; Score 428; DB 5; Length 3313;
 Best Local Similarity 100.0%; Pred. No. 2.7e-189;
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2757 AAGGCTGCAACACAGTGTAGCAGTGGCGGACACCATGTGCTTGAGCAGCATG 2816
 DB |||||
 DB 1218 AAGGCTGCAACACAGTGTAGCAGTGGCGGACACCATGTGCTTGAGCAGCATG 1277
 QY 2817 TGGAGATTGCACCGCGCAGCTGTAGTGTGGTGCAGCAACATGAAGCAGTGTGT 2876
 DB |||||
 DB 1278 TGGAGATTGCACCGCGCAGCTGTAGTGTGGTGCAGCAACATGAAGCAGTGTGT 1337
 QY 2877 GGACTCCCAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGTTACGAT 2936
 DB |||||
 DB 1338 GGACTCCCAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGTTACGAT 1397
 QY 2937 GAGCAGCTGCCCCCTGAAATTTTCAGGCTACTGTACCTGTAGTCTATGCTTGGAGCA 2996
 DB |||||
 DB 1398 GAGCAGCTGCCCCCTGAAATTTTCAGGCTACTGTACCTGTAGTCTATGCTTGGAGCA 1457
 QY 2997 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGGG 3056
 DB |||||
 DB 1458 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGGG 1517
 QY 3057 TTCCTATAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 3116
 DB |||||
 DB 1518 TTCCTATAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 1577
 QY 3117 GTCGTCAGTGTCTATAGAGGACAGCAGATACAACTGCTTTCATTCA 3176
 DB |||||


```
Db 1578 GCCCTGCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACACTGGTCTTTCAATCA 1637
QY 3177 CTGTCCAG 3184
Db 1638 CTGTCCAG 1645

RESULT 15
AA02409
ID AA02409 standard; cDNA; 668 BP.
XX AC
XX AA02409;
DT 18-JUL-2001 (first entry)
XX DE
XX Human secreted protein, cDNA #15.
XX Human; secreted protein; immunogen; antibody; diagnosis;
KW rheumatoid arthritis; hyperproliferative disorder; neoplasm;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; Alzheimer's disease; bacterial infection; viral infection;
KW fungal infection; corneal infection; wound healing; cell culture;
KW epithelial cell proliferation; skin ageing; transplantation;
KW tissue regeneration; chemotaxis; food additive; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200123546-A1.
XX PD
XX 05-APR-2001.
XX PF
XX 26-SEP-2000; 2000WO-US026323.
XX PR
XX 27-SEP-1999; 99US-0155805P.
XX PA
XX (HUMA-) HUMAN GENOME SCI INC.
XX PI
XX Rosen CA, Ruben S, Komatsoulis GA;
XX WPI; 2001-266150/27.
XX DR
XX P-PSDB; AAU01736, AAU01774.
XX PT
XX Nucleic acids encoding 37 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease
XX and diabetic retinopathy.
XX PS
XX Disclosure; Page 420; 494pp; English.
XX CC
XX The sequence encodes a human secreted protein of the invention. The
XX polynucleotides, polypeptides and antibodies raised against them are used
XX to prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, horses, cats, dogs, chickens or sheep. The
XX polynucleotides and antibodies are also used in diagnosing a pathological
XX condition or susceptibility to a pathological condition. The antibodies
XX can also be used in alleviating symptoms associated with the disorders
XX and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
XX immunosorbent assays (ELISA). Disorders which are diagnosed or treated
XX include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease,
XX infections caused by bacteria, viruses and fungi and ocular disorders
XX e.g. corneal infection. The polypeptides can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues and in chemotaxis. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities
XX SQ
XX Sequence 668 BP; 298 A; 113 C; 120 G; 137 T; 0 U; 0 Other;
Query Match 7.3%; Score 314; DB 4; Length 668;
Best Local Similarity 100.0%; Pred. No. 5.5e-136;
```

```
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3488 ATCAGTTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTG 3547
Db 9 ATCAGTTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTG 68
QY 3548 TGGCTACTCTGACGAACAACACAGGGATTGGACATGTTTCATCATGCTCCCAAGAAAT 3607
Db 69 TGGCTACTCTGACGAACAACACAGGGATTGGACATGTTTCATCATGCTCCCAAGAAAT 128
QY 3608 TCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGA 3667
Db 129 TCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGA 188
QY 3668 TGGCTGTGTTTCAAAAAACAACATTAAGGAGTACAAAGATGTTTCTTAATGAGAAGT 3727
Db 189 TGGCTGTGTTTCAAAAAACAACATTAAGGAGTACAAAGATGTTTCTTAATGAGAAGT 248
QY 3728 TTGATTTTGGCAACCCCAAAATATCATTCTCTTTGTTTATGTCAGTAAATTTTCACTGGC 3787
Db 249 TTGATTTTGGCAACCCCAAAATATCATTCTCTTTGTTTATGTCAGTAAATTTTCACTGGC 308
QY 3788 CCATCAAAATTCAG 3801
Db 309 CCATCAAAATTCAG 322

RESULT 16
AA072658
ID AA072658 standard; cDNA; 671 BP.
XX AC
XX AA072658;
DT 13-FEB-2002 (first entry)
XX DE
XX DNA encoding novel human diagnostic protein #8462.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200175067-A2.
XX PD
XX 11-OCT-2001.
XX PF
XX 30-MAR-2001; 2001WO-US008631.
XX PR
XX 31-MAR-2000; 2000US-00540217.
XX PR
XX 23-AUG-2000; 2000US-00649167.
XX PA
XX (HYSE-) HYSEQ INC.
XX PI
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR
XX P-PSDB; ABG08471.
XX PT
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX PS
XX Claim 1; SEQ ID NO 8462; 103pp; English.
XX CC
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
```

CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (ii) and its binding partners are useful in medical imaging
 CC of sites expressing (ii). (i) and (ii) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 671 BP; 207 A; 147 C; 148 G; 169 T; 0 U; 0 Other;
 Query Match 7.0%; Score 300; DB 5; Length 671;
 Best Local Similarity 100.0%; Pred. No. 2e-129;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3502 AGTCTATCCCGAAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTGAC 3561
 Db |||||
 QY 155 AGTCTATCCCGAAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTGAC 214
 Db |||||
 QY 3562 GAACAAAACAGGATTTGGACATGTTCAATGCTCCAAAGATTTCAACTCAACATC 3621
 Db |||||
 QY 215 GAACAAAACAGGATTTGGACATGTTCAATGCTCCAAAGATTTCAACTCAACATC 274
 Db |||||
 QY 3622 ACCTGGGTGCGAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCTGTGTTTCA 3681
 Db |||||
 QY 275 ACCTGGGTGCGAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCTGTGTTTCA 334
 Db |||||
 QY 3682 AAACCAACATTAAGGATGACAAAGATAGTTTCTCTATGAGAGTTTGATTTTGGCAAC 3741
 Db |||||
 QY 335 AAACCAACATTAAGGATGACAAAGATAGTTTCTCTATGAGAGTTTGATTTTGGCAAC 394
 Db |||||
 QY 3742 CACCCAAATATCATTCTTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAG 3801
 Db 395 CACCCAAATATCATTCTTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAG 454
 RESULT 17
 ID AAZ91916
 AC AAZ91916;
 XX
 XX AAZ91916;
 DT 08-JUN-2000 (first entry)
 DE Wild type (C57BL/6J) murine mahogany protein coding sequence.
 KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
 KW weight regulation; cell therapy; body weight disorder; cachexia;
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
 OS Mus sp.
 XX
 PN WO200005373-A2.
 XX
 PD 03-FEB-2000.
 XX
 PF 21-JUL-1999; 99WO-US016484.
 XX
 PR 21-JUL-1998; 98US-0093630P.
 PR 20-OCT-1998; 98US-0104978P.
 PR 05-FEB-1999; 99US-00245041.
 XX
 XX (MILL-) MILLENIUM PHARM INC.
 PA
 XX Moore K, Nagle DL;
 PI
 XX WPI; 2000-195103/17.
 DR

DR P-PSDB; AAY81803.
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and
 PT treatment of body weight disorders.
 XX
 PS Claim 1; Fig 2a; 188pp; English.
 XX
 CC This sequence represents a murine mahogany gene of the invention. The
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
 CC proteins (ii); (ii) as a source of antisense, ribozyme or triplex-forming
 CC therapeutics; (iii) as a source of diagnostic probes and primers for
 CC detecting expression of mg genes or mutations, regulatory defects, in
 CC this gene, or for isolation of related sequences; and (iv) in (cell-
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
 CC identify other (extra)cellular products involved in weight regulation,
 CC and to screen for agents that disrupt interaction between (ii) and other
 CC macromolecules. The Ab are used to detect abnormal levels (or function)
 CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to
 CC evaluate (ii)-expressing cells intended for cell therapy, and as
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
 CC mg polypeptide) are used to identify agents (A) that modulate mg
 CC activity. (A) are potentially useful for the treatment of body weight
 CC disorders, particularly obesity, cachexia or anorexia, or other
 CC conditions associated with the mg gene such as hyperpigmentation,
 CC hyperphagia and disorders that result in increased metabolic rate
 XX
 SQ Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;
 Query Match 6.8%; Score 291; DB 3; Length 8827;
 Best Local Similarity 100.0%; Pred. No. 3.1e-125;
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4000 GCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTC 4059
 Db 4266 GCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTC 4325
 QY 4060 AAACCCATTGCACTGGAGCCGTGTTTGGCAACAAAGCCGTCTCTCTCTGTTTGTG 4119
 Db 4326 AAACCCATTGCACTGGAGCCGTGTTTGGCAACAAAGCCGTCTCTCTCTGTTTGTG 4385
 QY 4120 AGSCTCCCTCGAGGCTGGTGGCATCCCTCTCTGGSCAGTCAGGTCTTCTGTGGCC 4179
 Db 4386 AGSCTCCCTCGAGGCTGGTGGCATCCCTCTCTGGSCAGTCAGGTCTTCTGTGGCC 4445
 QY 4180 AGCGCCCTGGTGACATTTCTCAGCAGATGCGCATAGTGTACAAGAGAGTCAGAGCC 4239
 Db 4446 AGCGCCCTGGTGACATTTCTCAGCAGATGCGCATAGTGTACAAGAGAGTCAGAGCC 4505
 QY 4240 GTGAGAAACCGAAGCAGCAGCCCTGACACAGCTGGGACCTGCATCTGA 4290
 Db 4506 GTGAGAAACCGAAGCAGCAGCCCTGACACAGCTGGGACCTGCATCTGA 4556
 RESULT 18
 ID ABK38928/c
 AC ABK38928 standard; cDNA; 246 BP.
 XX
 XX ABK38928;
 DT 21-MAY-2002 (first entry)
 DE cDNA encoding lung tumour protein clone R0127:F03.
 KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200204514-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 10-JUL-2001; 2001WO-US022058.

```
XX PR 11-JUL-2000; 2000US-00614124.
XX PR 29-AUG-2000; 2000US-00651563.
XX PR 08-SEP-2000; 2000US-00658824.
XX PR 26-SEP-2000; 2000US-00671325.
XX PR 06-OCT-2000; 2000US-00677419.
XX PR 30-OCT-2000; 2000US-00702705.
XX PR 13-DEC-2000; 2000US-00736457.
XX PR 03-MAY-2001; 2001US-00849626.
XX PA (CORI-) CORIYA CORP.
XX PR Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marxerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
PI Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX DR WPI; 2002-164634/21.
XX PR Novel polynucleotide encoding a lung tumor polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumor protein.
XX PS Claim 1; SEQ ID NO 966; 223pp; English.
XX CC The invention describes an isolated polynucleotide and polypeptide useful
CC for stimulating and/or expanding T cells specific for a tumour protein
CC containing the presence of a cancer in a patient. A composition
CC containing the polynucleotide and/or polypeptide is useful for treating a
CC lung cancer in a patient. The polypeptide is useful for removing tumour
CC cells from a biological sample. The polynucleotide is also useful as
CC probe or primer to detect the level of mRNA encoding a tumour protein.
CC This sequence encodes a lung tumour associated protein or protein
CC fragment, described in the method of the invention. Note: the sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
Query Match 5.7%; Score 246; DB 6; Length 246;
Best Local Similarity 100.0%; Freq. No. 3.5e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2515 CTGGAATAATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCATGGTGGC 2574
DB 246 CTGCGAATAATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCATGGTGGC 187
QY 2575 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCAATTAAGT 2634
DB 186 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCAATTAAGT 127
QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTCATCTGGATCTGTGGAATTTATCAGAA 2694
DB 126 TTACTACAGTGGATCCGCTCTGAGCCAGTCATCTGGATCTGTGGAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACTCAATGGTGTCTGT 2754
DB 66 CCCAGTACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACTCAATGGTGTCTGT 7
QY 2755 GAAAGG 2760
DB 6 GAAAGG 1
RESULT 19
ACAL1257/c
ID ACAL1257 standard; cDNA; 246 BP.
XX AC ACAL1257;
XX DT 05-JUN-2003 (first entry)
XX DE Human lung adenocarcinoma library cDNA SEQ ID 966.
XX
```

```
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX OS Homo sapiens.
XX PN US2002197669-A1.
XX PD 26-DEC-2002.
XX PF 03-MAY-2001; 2001US-00849626.
XX PR 13-DEC-2000; 2000US-00736457.
XX PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX DR WPI; 2003-352750/33.
XX PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
PT compositions, e.g. vaccines, for treating lung cancer.
XX PS Example 1; Page; 72pp; English.
XX CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
CC presenting cells that express the polynucleotide, such that T cells
CC proliferate, administering to the patient an effective amount of the
CC proliferated T cells, and thus inhibiting the development of a cancer in
CC the patient. The polynucleotide, protein and cells are useful in a
CC composition for stimulating an immune response in a patient, and for
CC treating a cancer in a patient (particularly lung cancer). The
CC oligonucleotide is useful for determining the presence of a cancer in a
CC patient. The protein and oligonucleotides are useful in pharmaceutical
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC or primer for nucleic acid hybridisation, and in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and proteins in tumour cells. An amplified portion of the
CC polynucleotide is useful for isolating a full-length gene from a suitable
CC library. The present sequence is a cDNA (full length, extended or
CC partial) isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
```

SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
Query Match 5.7%; Score 246; DB 7; Length 246;
Best Local Similarity 100.0%; Pred. No. 3.5e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2515 CTGCGAATAATGAGTCATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGAGTCATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGTGGC 187
QY 2575 CTTGCGAAGATCAATGTGTCCTACTGCTGGGAGATATGTCCTCCCAATTTACAATAGT 2634
Db 186 CTTGCGAAGATCAATGTGTCCTACTGCTGGGAGATATGTCCTCCCAATTTACAATAGT 127
QY 2635 TTACTACAGTGCATGCGCTGAGCCAGTATGCTGGATTCTGTGGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGCATGCGCTGAGCCAGTATGCTGGATTCTGTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1
RESULT 20
ACAO2443/C
ID ACA02443 standard; cDNA; 246 BP.
AC ACA02443;
DT 22-MAY-2003 (first entry)
DE Lung cancer therapy and diagnosis associated cDNA #932.
XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.
XX Homo sapiens.
XX US2002172952-A1.
XX 21-NOV-2002.
XX 10-JUL-2001; 2001US-00902941.
XX 30-JUN-1999; 99US-00346492.
XX 15-OCT-1999; 99US-00419356.
XX 17-DEC-1999; 99US-00466867.
XX 30-DEC-1999; 99US-00476300.
XX 06-MAR-2000; 2000US-00519642.
XX 22-MAR-2000; 2000US-00533077.
XX 10-APR-2000; 2000US-00546259.
XX 27-APR-2000; 2000US-00560406.
XX 05-JUN-2000; 2000US-00589184.
XX 11-JUL-2000; 2000US-00614124.
XX 29-AUG-2000; 2000US-00651563.
XX 08-SEP-2000; 2000US-00658824.
XX 26-SEP-2000; 2000US-00671325.
XX 06-OCT-2000; 2000US-00677419.
XX 30-OCT-2000; 2000US-00702705.
XX 13-DEC-2000; 2000US-00736457.
XX 03-MAY-2001; 2001US-00849626.
XX (CORI-) CORIXA CORP.
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;
PI Durham M, Carter D, Fanger GR, Vedvick TS, Bangor CS, McNabb A;
XX WPI; 2003-328427/31.
XX New polynucleotide, useful for preparing a composition for treating or

PT inhibiting development of cancer, e.g. lung cancer.
XX Example 1; SEQ ID NO 966; 82pp; English.
XX The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer. This
CC sequence represents a polynucleotide associated with the compositions and
CC methods for the therapy and diagnosis of lung cancer
XX
SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
Query Match 5.7%; Score 246; DB 7; Length 246;
Best Local Similarity 100.0%; Pred. No. 3.5e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2515 CTGCGAATAATGAGTCATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGAGTCATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGTGGC 187
QY 2575 CTTGCGAAGATCAATGTGTCCTACTGCTGGGAGATATGTCCTCCCAATTTACAATAGT 2634
Db 186 CTTGCGAAGATCAATGTGTCCTACTGCTGGGAGATATGTCCTCCCAATTTACAATAGT 127
QY 2635 TTACTACAGTGCATGCGCTGAGCCAGTATGCTGGATTCTGTGGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGCATGCGCTGAGCCAGTATGCTGGATTCTGTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1
RESULT 21
ABZ72040
ID ABZ72040 standard; DNA; 207433 BP.
XX AC ABZ72040;
XX DT 03-APR-2003 (first entry)
XX DE Gene 216 H194BAC1098L22 nucleotide sequence SEQ ID NO 5.
XX Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
KW antiinflammatory; gastrointestinal; gene therapy; vaccine; asthma;
KW obesity; inflammatory bowel disease; promoter; gene; ss.
XX Homo sapiens.
XX WO200178894-A2.
XX 25-OCT-2001.
XX 13-APR-2001; 2001WO-US012245.
XX 13-APR-2000; 2000US-00548797.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Keith T;
XX WPI; 2001-639428/73.
XX P-PSDB; ABR00926.
XX Isolated genes (Gene 216) from human chromosome 20p13-p12 and the
PT proteins they encode, useful for the prevention, diagnosis and treatment
PT of asthma, obesity and inflammatory bowel disease.
XX

```
PS Example 4; Fig 7; 520pp; English.
XX
CC The invention relates to isolated genes (Gene 216) from human chromosome
CC 20p13-p12 and the proteins they encode. The nucleic acids and proteins
CC may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate Gene 216 expression. For example, the
CC nucleic acids (or vectors) and proteins may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of gene 216 by expressing
CC inactive proteins or to supplement the patients own production of Gene
CC 216 proteins. Additionally, the nucleic acids may be used to produce the
CC secreted Gene 216 protein, by inserting the nucleic acids into a host
CC cell and culturing the cell to express the protein. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acid
CC sequences in samples and therefore which patients may be in need of
CC restorative therapy. The Gene 216 protein may also be used as antigens in
CC the production of antibodies against Gene 216 and in assays to identify
CC modulators of Gene 216 expression and activity. The anti-Gene 216
CC antibodies and antagonists may also be used to down regulate expression
CC and activity. The anti-Gene 216 antibodies may also be used as diagnostic
CC agents for detecting the presence of Gene 216 proteins in samples (e.g.
CC by enzyme linked immunosorbant assay or ELISA). Disorders that may be
CC prevented, diagnosed and/or treated by the above methods include, for
CC example asthma, obesity and inflammatory bowel disease. The present
CC sequence is that of the Gene 216 genomic nucleic acid sequence, promoter
CC or enhancer
XX
SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;
Query Match 5.5%; Score 234; DB 5; Length 207433;
Best Local Similarity 100.0%; Pred. No. 1.4e-98;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2951 CTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTGTTGAGCAACACCGCTGTGGCT 3010
DB 9813 CTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTGTTGAGCAACACCGCTGTGGCT 9872
QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 3070
DB 9873 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 9932
QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 3130
DB 9933 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 9992
QY 3131 CCAGCATGTGCTAGAGGACAGACATACAACTGGTCTTTTCATTCTACTGTCCAG 3184
DB 9993 CCAGCATGTGCTAGAGGACAGACATACAACTGGTCTTTTCATTCTACTGTCCAG 10046
RESULT 22
ABX74891
ID ABX74891 standard; DNA; 207433 BP.
XX
AC ABX74891;
XX
DT 07-APR-2003 (first entry)
XX
DE BAC1098L22 DNA sequence.
XX
KW Gene 216; antiasthmatic; antiinflammatory; ss; anorectic;
KW chromosome 20p13-p12; single nucleotide polymorphism; SNP; gene therapy;
KW respiratory disease; asthma; obesity; bronchial hyper-responsiveness;
KW chronic obstructive pulmonary disease;
KW adult respiratory distress syndrome; inflammatory bowel syndrome.
XX
OS Synthetic.
XX
PN WO200283077-A2.
XX
PD 24-OCT-2002.
XX
```

```
PF 15-APR-2002; 2002WO-US012063.
XX
PR 13-APR-2001; 2001US-00834597.
PR 13-APR-2001; 2001WO-US012245.
XX
PA (SCHE ) SCHERING CORP.
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Keith T, Little RD, Van Berdewegh P, Dupuis J, Del Mastro RG;
PI Simon J, Allen K, Pandit S;
XX
DR WPI; 2003-092960/08.
XX
PT New isolated gene 216 nucleic acids, useful for diagnosing, preventing or
PT treating a disorder, such as asthma, bronchial hyper-responsiveness,
PT chronic obstructive pulmonary disease, obesity or inflammatory bowel
PT syndrome.
XX
PS Example 6; Fig 7; 650pp; English.
XX
CC This invention relates to a novel isolated nucleic acid, gene 216,
CC identified from human chromosome 20p13-p12. The invention also discloses
CC regions of the 216 gene that contain single nucleotide polymorphisms
CC (SNP's) which may be used as markers for disease susceptibility or
CC severity. The nucleotides of the invention may have antiasthmatic,
CC antiinflammatory or anorectic activities and may be used in gene therapy.
CC The nucleic acids, antibodies or its fragments are useful for diagnosing,
CC preventing or treating a disorder, such as respiratory diseases (e.g.
CC asthma, bronchial hyper-responsiveness, chronic obstructive pulmonary
CC disease or adult respiratory distress syndrome), obesity, or inflammatory
CC bowel syndrome. The nucleic acids are also useful for identifying
CC increased susceptibility of a subject to the disorders mentioned. The
CC nucleic acids can also be used as primers and templates for the
CC recombinant production of disorder-associated peptides or polypeptides,
CC for chromosome and gene mapping, or for tissue distribution studies. The
CC present sequence represents a gene 216 cDNA sequence used in the scope of
CC the invention
XX
SQ Sequence 207433 BP; 52775 A; 51289 C; 51698 G; 51671 T; 0 U; 0 Other;
Query Match 5.5%; Score 234; DB 7; Length 207433;
Best Local Similarity 100.0%; Pred. No. 1.4e-98;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2951 CTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTGTTGAGCAACACCGCTGTGGCT 3010
DB 9813 CTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTGTTGAGCAACACCGCTGTGGCT 9872
QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 3070
DB 9873 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 9932
QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 3130
DB 9933 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 9992
QY 3131 CCAGCATGTGCTAGAGGACAGACATACAACTGGTCTTTTCATTCTACTGTCCAG 3184
DB 9993 CCAGCATGTGCTAGAGGACAGACATACAACTGGTCTTTTCATTCTACTGTCCAG 10046
RESULT 23
AAL14686/c
ID AAL14686 standard; cDNA; 481 BP.
XX
AC AAL14686;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 7143.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
```

OS Homo sapiens.
 XX WO200151628-A2.
 XX 19-JUL-2001.
 XX 10-JAN-2001; 2001WO-US000798.
 XX 14-JAN-2000; 2000US-0176077P.
 PR 14-MAR-2000; 2000US-0189167P.
 PR 24-MAR-2000; 2000US-0192099P.
 PR 29-MAR-2000; 2000US-0193480P.
 PR 15-MAY-2000; 2000US-0205230P.
 PR 09-JUN-2000; 2000US-0211315P.
 PR 25-JUL-2000; 2000US-0220534P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Lillie J, Xu Y, Wang Y, Steinmann K;
 XX WPI; 2001-451856/48.
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 PT Claim 1; Page 1292; 3695pp; English.
 XX The invention relates to human breast cancer expressed polynucleotides
 CC (AA10744-AAA26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity
 XX
 XX Sequence 481 BP; 106 A; 133 C; 126 G; 116 T; 0 U; 0 Other;
 SQ
 Query Match
 Best Local Similarity 5.3%; Score 228; DB 4; Length 481;
 Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2699 GTACTGGGGAGTGAAGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGCAAA 2758
 DB 279 GTACTGGGGAGTGAAGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGCAAA 220
 QY 2759 GGCCTGCAACACACAGTCTAAGCAGTGCAGGACCATGTCCTTGGACGACGATGTG 2818
 DB 219 GGCCTGCAACACACAGTCTAAGCAGTGCAGGACCATGTCCTTGGACGACGATGTG 160
 QY 2819 GAGATTGCACGAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGG 2878
 DB 159 GAGATTGCACGAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGG 100
 QY 2879 ACTCCAAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGTTACGATGA 2938
 DB 99 ACTCCAAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGTTACGATGA 40
 QY 2939 GCACCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACTT 2977
 DB 39 GCACCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACTT 1
 RESULT 24
 ID AAA45505
 XX AAA45505 standard; cDNA; 292 BP.
 AC AAA45505;
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human secreted expressed sequence tag SEQ ID NO:2080.
 XX
 KW Human, mouse; chicken; rat; secreted expressed sequence tag; sEST;
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
 KW antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
 KW autoimmune disorder; multiple sclerosis; allergic condition;
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KW central nervous system disorder; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; coagulation disorder;
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
 KW infection; depression; psoriasis; ss.
 XX
 OS Homo sapiens.
 XX WO2000021991-A1.
 PN 20-APR-2000.
 PD 15-OCT-1999; 99WO-US024206.
 PF 15-OCT-1998; 98US-0104436P.
 PR (GEMY) GENETICS INST INC.
 PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR;
 XX WPI; 2000-317938/27.
 XX Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (sESTs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders.
 XX
 XX Claim 1; Page 724; 803pp; English.
 XX AAA43426 to AAA45925 represent specifically claimed secreted expressed
 CC sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue
 CC sources. The sESTs can have a range of activities depending on the
 CC tissues they were isolated from. The activities include: chemotactic;
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;
 CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;
 CC antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antiulcer;
 CC osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;
 CC cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be
 CC used for gene therapy and in vaccines. The sESTs are useful as probes for
 CC the identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention
 XX
 SQ Sequence 292 BP; 77 A; 58 C; 71 G; 86 T; 0 U; 0 Other;
 Query Match
 Best Local Similarity 5.3%; Score 227; DB 3; Length 292;
 Matches 277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 717 TCGATTATATTTACTTACAGTTTGTATGTGTCCTCAATTAATCTCTCAGCCGAGGAGA 776
 DB 15 TCGATTATATTTACTTACAGTTTGTATGTGTCCTCAATTAATCTCTCAGCCGAGGAGA 74
 QY 777 GTGTGAAGATCATGTAATAGCAGCAACTGTTGAATGTGAATGTTCTGAAACTGGAAGG 836

Db 75 GTGTAGATCAGTAAATAGCAGGACTACTGTTGATGTGATGTTCTGAAACTGGAAGG 134

Qy 837 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCCCTCATCGAGGATCTG 896

Db 135 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCCCTCATCGAGGATCTG 194

Qy 897 CAATTCAGTGATGTACAGGATGCTCTCTCTCTCACTGTGCGAGGTCCTCGATGTTTC 956

Db 195 CAATTCAGTGATGTACAGGATGCTCTCTCTCTCACTGTGCGAGGTCCTCGATGTTTC 254

Qy 957 AGTTCCTGTACCACTAACCAGTCAATTTGGACTCGAG 994

Db 255 AGTTCCTGTACCACTAACCAGTCAATTTGGACTCGAG 292

RESULT 25

AAI43425

ID AAI43425 standard; DNA; 531 BP.

XX AC AAI43425;

XX DT 17-OCT-2001 (first entry)

XX DE Probe #12111 used to measure gene expression in human placenta sample.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX PS gene expression in human placenta.

XX PS Claim 25; SEQ ID NO 12111; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP).

XX CC The present sequence is one such probe. The probes are useful for

XX CC producing a microarray for predicting, measuring and displaying gene

XX CC expression in samples derived from human placenta. The probes are useful

XX CC for antenatal diagnosis of human genetic disorders

XX SQ Sequence 531 BP; 145 A; 82 C; 107 G; 197 T; 0 U; 0 Other;

Query Match 4.9%; Score 210; DB 4; Length 531;

Best Local Similarity 100.0%; Pred. No. 2.4e-87;

Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1238 CAATCGGAATGTGACCAATGATGAGTGTGAGTTTTCATTCATTAAGTCAAGTGGTGT 1297

Db 184 CAATCGGAATGTGACCAATGATGAGTGTGAGTTTTCATTCATTAAGTCAAGTGGTGT 243

Qy 1298 TGTTGACCCCTAAGGCAAGGACGATGATGAGTGGTGGCTCTGACACATTTGTTA 1357

Db 244 TGTTGACCCCTAAGGCAAGGACGATGATGAGTGGTGGCTCTGACACATTTGTTA 303

Qy 1358 CACTGAAGAATGGCGAGTGTGATGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGAT 1417

Db 304 CACTGAAGAATGGCGAGTGTGATGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGAT 363

Qy 1418 ATATAAGCAATGTGAGGAATATGATTTGG 1447

Db 364 ATATAAGCAATGTGAGGAATATGATTTGG 393

RESULT 26

ABS11523

ID ABS11523 standard; DNA; 531 BP.

XX AC ABS11523;

XX DT 19-AUG-2002 (first entry)

XX DE Human genome-derived single exon probe from lung SEQ ID No 11514.

XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX KW chronic obstructive pulmonary disease; interstitial lung disease;

XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;

XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX KW primary ciliary dyskinesia; pulmonary hypertension;

XX KW hyaline membrane disease.

XX OS Homo sapiens.

XX PN WO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001WO-US000665.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to

XX PS measure gene expression in human lung samples.

XX PS Claim 1; SEQ ID NO 11514; 634pp; English.

XX CC The invention relates to a spatially-addressable set of single exon

XX CC nucleic acid probes for measuring gene expression in a sample derived

XX CC from human lung comprising single exon nucleic acid probes having one of

XX CC 12614 nucleic acid sequences mentioned in the specification, or their

XX CC complements or the 12387 open reading frames derived from the 12614

XX CC probes. Also included are a microarray comprising the novel set of probes

XX CC ; the novel set of probes which hybridise at high stringency to a nucleic

XX CC acid expressed in the human lung; measuring gene expression in a nucleic

XX CC derived from human lung, comprising (a) contacting the array with a

XX CC collection of detectably labeled nucleic acids derived from human lung

XX CC mRNA, and (b) measuring the label detectably bound to each probe of the

XX CC array; identifying exons in a eukaryotic genome, comprising (a)

XX CC algorithmically predicting at least one exon from genomic sequences of

XX CC the eukaryote; and (b) detecting specific hybridisation of detectably

XX CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 531 BP; 145 A; 82 C; 107 G; 197 T; 0 U; 0 Other;

Query Match 4.9%; Score 210; DB 6; Length 531;
 Best Local Similarity 100.0%; Pred. No. 2.4e-87;
 Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1238 CAACCTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAAGTGCATGGGTG 1297
 Db 184 CAACCTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAAGTGCATGGGTG 243
 QY 1298 TGTGACCCCTAAGGCAAGAGCAGTATGAGTGGTGGGCACTCTGCACACATTGTTA 1357
 Db 244 TGTGACCCCTAAGGCAAGAGCAGTATGAGTGGTGGGCACTCTGCACACATTGTTA 303
 QY 1358 CACTGACGAATGGCGAGTGGTGCATGCTGGTGCATCTTTGGTGCATGCTCCCTCTATGGAT 1417
 Db 304 CACTGACGAATGGCGAGTGGTGCATGCTGGTGCATCTTTGGTGCATGCTCCCTCTATGGAT 363
 QY 1418 ATATAAGCAATGTGCAGCAATATGATTGG 1447
 Db 364 ATATAAGCAATGTGCAGCAATATGATTGG 393

RESULT 27
 AA156405
 ID AA156405 standard; DNA; 193 BP.
 AC
 XX
 XX
 XX

DT 17-OCT-2001 (first entry)

DE Probe #25091 used to measure gene expression in human placenta sample.

KW Probe; microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

XX Homo sapiens.

OS WO200157272-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US000663.

PF 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX
 PA (MOLS-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX MPI; 2001-488897/53.
 DR
 DR
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human placenta.
 PS Claim 25; SEQ ID NO 25091; 654pp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes (SENPs).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders
 XX
 SQ Sequence 193 BP; 50 A; 36 C; 52 G; 55 T; 0 U; 0 Other;

Query Match 4.5%; Score 193; DB 4; Length 193;
 Best Local Similarity 100.0%; Pred. No. 2.1e-79;
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1245 GAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAAGTGCATGGGTGTTGAC 1304
 Db 1 GAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAAGTGCATGGGTGTTGAC 60
 QY 1305 CCTTAAAGCAAGAGCAGTATGAGTGGTGGGCACTCTGCACACATTGTACCTGAA 1364
 Db 61 CCTTAAAGCAAGAGCAGTATGAGTGGTGGGCACTCTGCACACATTGTACCTGAA 120
 QY 1365 GAATGGCCGAGTGGTGCATGCTGGTGCATCTTTGGTGCATGCTCCCTCTATGGATATAAG 1424
 Db 121 GAATGGCCGAGTGGTGCATGCTGGTGCATCTTTGGTGCATGCTCCCTCTATGGATATAAG 180
 QY 1425 CAATGTGCAGGAA 1437
 Db 181 CAATGTGCAGGAA 193

RESULT 28

ABS23881
 ID ABS23881 standard; DNA; 193 BP.
 AC
 XX
 XX
 XX
 XX
 XX
 XX

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe ORF from lung SEQ ID No 23872.

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

Db 62 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

QY 2725 TG 2726
||
2 TG 1

Db 2 TG 1

RESULT 30
ABR69223/c
ID ABA69223 standard; DNA; 182 BP.
XX AC ABA69223;
XX AC ABA69223;
XX DT 01-FEB-2002 (first entry)
XX DE Human foetal liver single exon nucleic acid probe #17528.
XX DE Human foetal liver; gene expression; single exon nucleic acid probe; ss.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX EN 09-AUG-2001.
XX PD 30-JAN-2001; 2001WO-US000669.
XX PF 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-483447/52.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human fetal liver.
XX PS Claim 4; SEQ ID NO 17528; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human fetal liver. The
XX CC present sequence is a single exon nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences

QY 2545 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604
Db 182 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 123

QY 2605 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGGTCTGAGCCCACT 2664
Db 122 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGGTCTGAGCCCACT 63

QY 2665 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724
Db 62 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

QY 2725 TG 2726
||
2 TG 1

Db 2 TG 1

RESULT 31
AAI49392/c
ID AAI49392 standard; DNA; 182 BP.
XX AC AAI49392;
XX AC AAI49392;
XX DT 17-OCT-2001 (first entry)
XX DE Probe #18078 used to measure gene expression in human placenta sample.
XX DE Probe; microarray; human; placenta; antenatal diagnosis;
XX KW Genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX EN 09-AUG-2001.
XX PD 30-JAN-2001; 2001WO-US000663.
XX PF 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-488897/53.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 25; SEQ ID NO 18078; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders

QY Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Query Match 4.2%; Score 182; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.9e-74;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604
Db 182 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 123

QY 2605 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGGTCTGAGCCCACT 2664
Db 122 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGGTCTGAGCCCACT 63

QY 2665 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724
Db 62 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

QY 2725 TG 2726
||
2 TG 1

Db 2 TG 1

RESULT 32
ABA51213/c
ID ABA51213 standard; DNA; 182 BP.
XX
AC ABA51213;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #9908.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 4; SEQ ID NO 9908; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
Query Match 4.2%; Score 182; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.9e-74;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2545 TCCAGCTCACCTTAACCCCATGGGTTCGGAGATCAATGTCTCTACTGGTGC 2604
DB 182 TCCAGCTCACCTTAACCCCATGGGTTCGGAGATCAATGTCTCTACTGGTGC 123
QY 2605 TGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATCGCTCTGAGCCCACT 2664

DB 122 TGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATCGCTCTGAGCCCACT 63
QY 2665 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACC 2724
DB 62 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACC 3
QY 2725 TG 2726
DB 2 TG 1
RESULT 33
ABA36140/c
ID ABA36140 standard; DNA; 182 BP.
XX
AC ABA36140;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #14606 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 14606; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
Query Match 4.2%; Score 182; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.9e-74;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2545 TCCAGCTCACCTTAACCCCATGGGTTCGGAGATCAATGTCTCTACTGGTGC 2604
DB 182 TCCAGCTCACCTTAACCCCATGGGTTCGGAGATCAATGTCTCTACTGGTGC 123

QY 2605 TGGGAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 2664
 Db 122 TGGGAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 63
 QY 2665 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724
 Db 62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
 QY 2725 TG 2726
 Db 2 TG 1

RESULT 34
 AAK43324/c
 ID AAK43324 standard; DNA; 182 BP.
 XX
 AC AAK43324;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed single exon probe SEQ ID NO: 17881.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000668.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-489900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 XX
 PS Example 4; SEQ ID NO 17881; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention
 XX
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
 Query Match 4.2%; Score 182; DB 4; Length 182;
 Best Local Similarity 100.0%; Pred. No. 2.9e-74;
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2545 TCCAGCTCAGCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604
 Db 182 TCCAGCTCAGCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 123
 QY 2605 TGGGAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 2664

Db 122 TGGGAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 63
 QY 2665 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724
 Db 62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
 QY 2725 TG 2726
 Db 2 TG 1

RESULT 35
 AAK17517/c
 ID AAK17517 standard; DNA; 182 BP.
 XX
 AC AAK17517;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 17508.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000667.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO 17508; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention
 XX
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
 Query Match 4.2%; Score 182; DB 4; Length 182;
 Best Local Similarity 100.0%; Pred. No. 2.9e-74;
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2545 TCCAGCTCAGCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604
 Db 182 TCCAGCTCAGCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 123
 QY 2605 TGGGAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 2664
 Db 122 TGGGAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 63

Qy 2665 GATGCTGATTCTCTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724
Db 62 GATGCTGATTCTCTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
Qy 2725 TG 2726
Db 2 TG 1
RESULT 36
ABS42949/C
ID ABS42949 standard; DNA; 182 BP.
XX ABS42949;
XX
XX 25-FEB-2003 (first entry)
XX Human liver single exon probe, SEQ ID NO 17939.
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX Homo sapiens.
XX WO200157273-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000664.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX Claim 4; SEQ ID NO 17939; 658pp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
Query Match 4.2%; Score 182; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.9e-74;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2545 TCCAAAGTCACTTAAACCCAGTGGTGGCCCTTCGGAAGATCAATGTGCTACTGGTGC 2604

Db 182 TCCAAAGTCACTTAAACCCAGTGGTGGCCCTTCGGAAGATCAATGTGCTACTGGTGC 123
Qy 2605 TGGGAAGATATCTCCCAATTTTACAAATAGTTTACTACAGTGGATGCCGCTTGAGCCCACT 2664
Db 122 TGGGAAGATATCTCCCAATTTTACAAATAGTTTACTACAGTGGATGCCGCTTGAGCCCACT 63
Qy 2665 GATGCTGATTCTCTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724
Db 62 GATGCTGATTCTCTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
Qy 2725 TG 2726
Db 2 TG 1
RESULT 37
AAI09674/C
ID AAI09674 standard; DNA; 182 BP.
XX AAI09674;
XX
XX 09-OCT-2001 (first entry)
XX Probe #9665 used to measure gene expression in human breast sample.
XX Probe; human; breast disease; breast cancer; development disorder; ss;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX Homo sapiens.
XX WO200157270-A2.
XX 09-AUG-2001.
XX 29-JAN-2001; 2001WO-US000661.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX Novel single exon nucleic acid probe used to measuring gene expression in
XX a human breast.
XX Claim 25; SEQ ID NO 9665; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes.
XX The present sequence is one such probe. The probes are useful for
XX measuring human gene expression in a human breast sample, where the probe
XX hybridises at high stringency to a nucleic acid expressed in the human
XX breast. The probes are useful for predicting, diagnosing, grading,
XX staging, monitoring and prognosing diseases of the human breast,
XX particularly those diseases with polygenic aetiology. The diseases
XX include: breast cancer, disorders of development, inflammatory diseases
XX of the breast, fibrocystic changes, proliferative breast disease and non-
XX carcinoma tumours. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
Query Match 4.2%; Score 182; DB 5; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.9e-74;

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGTGGCTTCCGAGATCAATGTCTCTACTGTGC 2604
 Db 182 TCCAGCTCACCTTAACCCCATGGTGGCTTCCGAGATCAATGTCTCTACTGTGC 123

QY 2605 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCCGCTGAGCCAGT 2664
 Db 122 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCCGCTGAGCCAGT 63

QY 2665 GATGCTGATTCGTGGAAATTTATCAGAACCCAGTACTCGGACATGAGGCTGCAACC 2724
 Db 62 GATGCTGATTCGTGGAAATTTATCAGAACCCAGTACTCGGACATGAGGCTGCAACC 3

QY 2725 TG 2726
 Db 2 TG 1

RESULT 39
 ABS17404/c
 ID ABS17404 standard; DNA; 182 BP.
 AC ABS17404;
 XX
 XX
 DT 19-AUG-2002 (first entry)
 DE Human genome-derived single exon probe ORF from lung SEQ ID No 17395.
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX Homo sapiens.
 OS
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 FA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 4; SEQ ID NO 17395; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC ; the novel set of probes which hybridise at high stringency to a nucleic

acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Query Match 4.2%; Score 182; DB 6; Length 182;
 Best Local Similarity 100.0%; Pred. No. 2.9e-74;
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGTGGCTTCCGAGATCAATGTCTCTACTGTGC 2604
 Db 182 TCCAGCTCACCTTAACCCCATGGTGGCTTCCGAGATCAATGTCTCTACTGTGC 123

QY 2605 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCCGCTGAGCCAGT 2664
 Db 122 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCCGCTGAGCCAGT 63

QY 2665 GATGCTGATTCGTGGAAATTTATCAGAACCCAGTACTCGGACATGAGGCTGCAACC 2724
 Db 62 GATGCTGATTCGTGGAAATTTATCAGAACCCAGTACTCGGACATGAGGCTGCAACC 3

QY 2725 TG 2726
 Db 2 TG 1

RESULT 39
 AA114910/c
 ID AA114910 standard; DNA; 375 BP.
 XX
 AC AA114910;
 XX
 XX 12-OCT-2001 (first entry)
 DT
 DE Probe #4843 for gene expression analysis in human cervical cell sample.
 XX
 XX Probe; human; microarray; gene expression; cervical epithelial cell;
 XX cervical cancer; ss.
 KW
 KW Homo sapiens.
 OS
 XX WO200157278-A2.
 PN
 XX 09-AUG-2001.
 PD


```

XX 30-JAN-2001; 2001WO-US000670.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 04-OCT-2000; 2000US-0236359P.
XX 04-FEB-2000; 2000US-0180312P.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-489901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX Claim 25; SEQ ID NO 4843; 487pp; English.
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human Hela cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging of
XX diseases of the cervix, notably cervical cancer. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
XX
XX Query Match 4.0%; Score 173; DB 4; Length 375;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-70;
XX Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2592 GTCCTACTGGTCTGGGAAGATATCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651
XX DB 375 GTCTTACTGGTCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316
XX
XX QY 2652 GTCTGAGCCCAAGTATGCTGGATTCTGGAATTTATCAGAACCCAGTACTCGGGGACT 2711
XX DB 315 GTCTGAGCCCAAGTATGCTGGATTCTGGAATTTATCAGAACCCAGTACTCGGGGACT 256
XX
XX QY 2712 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
XX DB 255 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
XX
XX RESULT 40
XX ABA56648/c
XX ID ABA56648 standard; DNA; 375 BP.
XX AC ABA56648;
XX
XX DT 01-FEB-2002 (first entry)
XX
XX DE Human foetal liver single exon nucleic acid probe #4953.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000669.
XX
XX QY 04-FEB-2000; 2000US-0180312P.

```

```

PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000US-0236359P.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX Claim 1; SEQ ID NO 4953; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human fetal liver. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
XX
XX Query Match 4.0%; Score 173; DB 4; Length 375;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-70;
XX Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2592 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651
XX DB 375 GTCTTACTGGTCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316
XX
XX QY 2652 GTCTGAGCCCAAGTATGCTGGATTCTGGAATTTATCAGAACCCAGTACTCGGGGACT 2711
XX DB 315 GTCTGAGCCCAAGTATGCTGGATTCTGGAATTTATCAGAACCCAGTACTCGGGGACT 256
XX
XX QY 2712 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
XX DB 255 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
XX
XX RESULT 41
XX AAI36262/c
XX ID AAI36262 standard; DNA; 375 BP.
XX AC AAI36262;
XX
XX DT 17-OCT-2001 (first entry)
XX
XX DE Probe #4948 used to measure gene expression in human placenta sample.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157272-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000663.
XX
XX QY 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.

```

PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488997/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 4948; 554pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Query Match 4.0%; Score 173; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2651
DB 375 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2711
DB 315 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 256
QY 2712 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
DB 255 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
RESULT 42
ABA46100/c
ID ABA46100 standard; DNA; 375 BP.
XX
AC ABA46100;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #4795.
XX
DE Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX

PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 4795; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Query Match 4.0%; Score 173; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2651
DB 375 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2711
DB 315 GTCCTACTGGTCTGGGAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 256
QY 2712 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
DB 255 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
RESULT 43
ABA26254/c
ID ABA26254 standard; DNA; 375 BP.
XX
AC ABA26254;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #4720 for gene expression analysis in human heart cell sample.
XX
DE Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX


```

PT brains.
XX Example 4; SEQ ID NO 4776; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
XX Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
    Query Match      4.0%; Score 173; DB 4; Length 375;
    Best Local Similarity 100.0%; Pred. No. 4.7e-70;
    Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651
DB 375 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCCTAGCCAGTGTGCTGGATTCCTGGAAATTTATCAGAACCCAGTACTCGGGACT 2711
DB 315 GTCCTAGCCAGTGTGCTGGATTCCTGGAAATTTATCAGAACCCAGTACTCGGGACT 256
QY 2712 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 2764
DB 255 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 203

RESULT 46
ABS29950/c
ID ABS29950 standard; DNA; 375 BP.
XX
AC ABS29950;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID NO 4940.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024253.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
XX
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PI WPI; 2001-488898/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 1; SEQ ID NO 4940; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SEN) (I) for

```

```

CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (i) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
    Query Match      4.0%; Score 173; DB 4; Length 375;
    Best Local Similarity 100.0%; Pred. No. 4.7e-70;
    Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651
DB 375 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCCTAGCCAGTGTGCTGGATTCCTGGAAATTTATCAGAACCCAGTACTCGGGACT 2711
DB 315 GTCCTAGCCAGTGTGCTGGATTCCTGGAAATTTATCAGAACCCAGTACTCGGGACT 256
QY 2712 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 2764
DB 255 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 203

RESULT 47
AAI04689/c
ID AAI04689 standard; DNA; 375 BP.
XX
AC AAI04689;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #4680 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
XX
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PI WPI; 2001-476286/51.
XX
DR Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 4680; 322pp; English.
XX
XX

```

```
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Query Match 4.0%; Score 173; DB 5; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCTACTGGTCTGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATGCC 2651
Db GTCTACTGGTCTGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTGAGCCCGAGTGGATCTGGAATTTATCAGAACCCAGTACTCGGGACT 2711
Db GTCTGAGCCCGAGTGGATCTGGAATTTATCAGAACCCAGTACTCGGGACT 256
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
Db GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
RESULT 49
ABS04892/c
ID ABS04892 standard; DNA; 375 BP.
AC ABS04892;
XX
XX 19-AUG-2002 (first entry)
DE Human genome-derived single exon probe from lung SEQ ID No 4883.
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 1; SEQ ID NO 4883; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridise at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array, identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene expression
XX analysis, and for identifying exons in a gene, particularly using human
XX lung derived mRNA and for the study of lung diseases such as asthma, lung
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Heremansky-
XX Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX present sequence is a single exon probe of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Query Match 4.0%; Score 173; DB 6; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCTACTGGTCTGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATGCC 2651
Db GTCTACTGGTCTGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTGAGCCCGAGTGGATCTGGAATTTATCAGAACCCAGTACTCGGGACT 2711
Db GTCTGAGCCCGAGTGGATCTGGAATTTATCAGAACCCAGTACTCGGGACT 256
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
Db GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
RESULT 49
AAS69062
ID AAS69062 standard; cDNA; 2267 BP.
XX
XX AAS69062;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #4866.
```

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX P-PSDB; ABGC4875.
 XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.
 XX Claim 1; SEQ ID NO 4866; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (II) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Query Match 3.4%; Score 148; DB 5; Length 2267;
 XX Best Local Similarity 100.0%; Pred. No. 2.3e-58;
 XX Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX 3417 ATGTCAGGTAGAAATCGATACAGAGAACCTCTCAGAGGACATGTTATTACTCT 3476
 XX 162 ATGTCAGGTAGAAATCGATACAGAGAACCTCTCAGAGGACATGTTATTACTCT 221
 XX 3477 TCTTTATGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGC 3536
 XX 222 TCTTTATGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGC 281
 XX 3537 TATCAATTTTGGTCTACTCTCAGCAA 3564
 XX 282 TATCAATTTTGGTCTACTCTCAGCAA 309

XX RESULT 50
 XX AAI13789/c
 XX ID AAI13789 standard; DNA; 402 BP.

XX AC AAI13789;
 XX AC 12-OCT-2001 (first entry)
 XX DE Probe #3722 for gene expression analysis in human cervical cell sample.
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer; ss.
 XX OS Homo sapiens.
 XX WO200157278-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000670.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000US-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells.
 XX Claim 25; SEQ ID NO 3722; 487pp; English.
 XX The present invention relates to human single exon nucleic acid probes (SENPs). The present sequence is one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Query Match 3.3%; Score 141; DB 4; Length 402;
 XX Best Local Similarity 100.0%; Pred. No. 4.3e-55;
 XX Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX 3182 CAGCTTGCCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTG 3241
 XX 391 CAGCTTGCCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTG 332
 XX 3242 AGAAGCTTGCCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTG 3301
 XX 331 AGAAGCTTGCCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTG 272
 XX 3302 CCAATGGAGGAAATGTCAGC 3322
 XX 271 CCAATGGAGGAAATGTCAGC 251

XX RESULT 51
 XX ABA55499/c
 XX ID ABA55499 standard; DNA; 402 BP.
 XX AC ABA55499;
 XX AC ABA55499;

```
DT 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #3804.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human foetal liver.
XX
XX Claim 1; SEQ ID NO 3804; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human foetal liver. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
XX
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241
DB 391 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 332
QY 3242 AGAACCCTGACCAAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCA 3301
DB 331 AGAACCCTGACCAAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCA 272
QY 3302 CCAATGGAGGAAATGTCAGC 3322
DB 271 CCAATGGAGGAAATGTCAGC 251
RESULT 52
AAI35149/C
ID AAI35149 standard; DNA; 402 BP.
XX
XX AAI35149;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #3835 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
```

```
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
XX Claim 25; SEQ ID NO 3835; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders
XX
XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
XX
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241
DB 391 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 332
QY 3242 AGAACCCTGACCAAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCA 3301
DB 331 AGAACCCTGACCAAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCA 272
QY 3302 CCAATGGAGGAAATGTCAGC 3322
DB 271 CCAATGGAGGAAATGTCAGC 251
RESULT 53
ABA45027/C
ID ABA45027 standard; DNA; 402 BP.
XX
XX ABA45027;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #3722.
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000662.
```



```
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX PS Claim 1; SEQ ID NO 3722; 327pp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the invention. Note: The sequence data for this
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCCCAATGCAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241
DB 391 CAGCTTGCCCAATGCAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332
QY 3242 AGAACCTGACCAAGCAGCAAGCACTGCGAGACTGCGATATCTGGCTTCTACGGTGATCCCA 3301
DB 331 AGAACCTGACCAAGCAGCAAGCACTGCGAGACTGCGATATCTGGCTTCTACGGTGATCCCA 272
QY 3302 CCAATGGAGGAAATGTGAGC 3322
DB 271 CCAATGGAGGAAATGTGAGC 251
RESULT 54
ABA25215/c
ID ABA25215 standard; DNA; 402 BP.
XX AC ABA25215;
XX DT 23-JAN-2002 (first entry)
XX DE Probe #3681 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX XX
```

```
OS XX Homo sapiens.
XX PN WO200157274-A2.
XX XX
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX PS Claim 1; SEQ ID NO 3681; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCCCAATGCAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241
DB 391 CAGCTTGCCCAATGCAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332
QY 3242 AGAACCTGACCAAGCAGCAAGCACTGCGAGACTGCGATATCTGGCTTCTACGGTGATCCCA 3301
DB 331 AGAACCTGACCAAGCAGCAAGCACTGCGAGACTGCGATATCTGGCTTCTACGGTGATCCCA 272
QY 3302 CCAATGGAGGAAATGTGAGC 3322
DB 271 CCAATGGAGGAAATGTGAGC 251
RESULT 55
AAK29197/c
ID AAK29197 standard; DNA; 402 BP.
XX AC AAK29197;
XX XX
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 3754.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX XX
```

```

PN WO200157276-A2.
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US000668.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
PD 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX Example 4; SEQ ID NO 3754; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
SQ
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241
DB 391 CAGCTTGCATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 332
QY 3242 AGAAGCTGACCAAGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 3301
DB 331 AGAAGCTGACCAAGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 272
QY 3302 CCAATGGAGGAAATGTCAGC 3322
DB 271 CCAATGGAGGAAATGTCAGC 251
RESULT 56
AAK03733/C
ID AAK03733 standard; DNA; 402 BP.
XX AAK03733;
XX AC
XX AAK03733;
XX 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe SEQ ID NO: 3724.
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX Homo sapiens.
XX WO200157275-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000667.
XX
PN WO200157276-A2.
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US000668.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
PD 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX Example 4; SEQ ID NO 3754; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
SQ
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241
DB 391 CAGCTTGCATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 332
QY 3242 AGAAGCTGACCAAGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 3301
DB 331 AGAAGCTGACCAAGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 272
QY 3302 CCAATGGAGGAAATGTCAGC 3322
DB 271 CCAATGGAGGAAATGTCAGC 251
RESULT 57
ABS28814/C
ID ABS28814 standard; DNA; 402 BP.
XX ABS28814;
XX AC
XX ABS28814;
XX 25-FEB-2003 (first entry)
XX Human liver single exon probe, SEQ ID No 3804.
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX Homo sapiens.
XX WO200157273-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000664.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX

```

PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488998/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 XX Claim 1; SEQ ID NO 3804; 658pp; English.
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridizes at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABS25011-ABS51005 represent human
 CC liver single exon nucleic acid probes of the invention. Note: The
 CC sequence information for this patent does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
 Query Match 3.3%; Score 141; DB 4; Length 402;
 Best Local Similarity 100.0%; Pred. No. 4.3e-55;
 Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3182 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241
 DB |||||
 391 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332
 QY 3242 AGAACCTGACCAAGCAGCAAGCACTGCGAGCTGATATCTGGCTTCTACGGTGATCCCA 3301
 DB |||||
 331 AGAACCTGACCAAGCAGCAAGCACTGCGAGCTGATATCTGGCTTCTACGGTGATCCCA 272
 QY 3302 CCAATGGAGGAAATGTGAGC 3322
 DB |||||
 271 CCAATGGAGGAAATGTGAGC 251
 RESULT 58
 AA103665/c
 ID AA103665 standard; DNA; 402 BP.
 XX
 AC AA103665;
 XX
 DT 09-OCT-2001 (first entry)
 XX
 DE Probe #3656 used to measure gene expression in human breast sample.
 XX Probe; human; breast disease; breast cancer; development disorder; ss;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX Homo sapiens.
 OS
 XX WO200157270-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 29-JAN-2001; 2001WO-US000661.
 XX
 XX 04-FEB-2000; 2000US-0180312P.
 PR
 PR 26-MAY-2000; 2000US-0207458P.
 PR 30-JUN-2000; 2000US-0060840B.
 PR 03-AUG-2000; 2000US-00632365.

PR 21-SEP-2000; 2000US-0234587P.
 PR 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.
 XX Novel single exon nucleic acid probe used to measuring gene expression in
 PT a human breast.
 XX Claim 25; SEQ ID NO 3656; 322pp; English.
 XX The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridizes at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and non-
 CC carcinoma tumours. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
 Query Match 3.3%; Score 141; DB 5; Length 402;
 Best Local Similarity 100.0%; Pred. No. 4.3e-55;
 Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3182 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241
 DB |||||
 391 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332
 QY 3242 AGAACCTGACCAAGCAGCAAGCACTGCGAGCTGATATCTGGCTTCTACGGTGATCCCA 3301
 DB |||||
 331 AGAACCTGACCAAGCAGCAAGCACTGCGAGCTGATATCTGGCTTCTACGGTGATCCCA 272
 QY 3302 CCAATGGAGGAAATGTGAGC 3322
 DB |||||
 271 CCAATGGAGGAAATGTGAGC 251
 RESULT 59
 ABS03749/c
 ID ABS03749 standard; DNA; 402 BP.
 XX
 AC ABS03749;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 3740.
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 XX WO200186003-A2.
 PN
 XX 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX Claim 1; SEQ ID NO 3740; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridization of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridization to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Query Match 3.3%; Score 141; DB 6; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3182 CAGCTTGCCAAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 3241
DB 391 CAGCTTGCCAAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 332
QY 3242 AGAAGCTGACACAGCAAGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 3301
DB 331 AGAAGCTGACACAGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 272

QY 3302 CCAATGGAGGAAATGTCAGC 3322
DB 271 CCAATGGAGGAAATGTCAGC 251

RESULT 60
AAI23009/c
ID AAI23009 standard; DNA; 137 BP.
XX
AC AAI23009;
XX
XX 12-OCT-2001 (first entry)
XX
DE Probe #12942 for gene expression analysis in human cervical cell sample.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WC200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
XX Claim 25; SEQ ID NO 12942; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Query Match 3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3185 CTTGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 3244
DB 137 CTTGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 78
QY 3245 ACCTGACACAGGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 3304
DB 77 ACCTGACACAGGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 18
QY 3305 ATGAGGGAATGTCAG 3321

```

Db      17 ATGGAGGGAATGTCTAG 1
RESULT 61
ABR68099/c
ID   ABA68099 standard; DNA; 137 BP.
XX
AC   ABA68099;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16404.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
XX
PR   26-MAY-2000; 2000US-0207456P.
XX
PR   30-JUN-2000; 2000US-00608408.
XX
PR   03-AUG-2000; 2000US-00632366.
XX
PR   21-SEP-2000; 2000US-0234687P.
XX
PR   27-SEP-2000; 2000US-0236359P.
XX
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
XX   gene expression in human foetal liver.
XX
PS   Claim 4; SEQ ID NO 16404; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
XX   human gene expression in a sample derived from human foetal liver. The
XX   single exon nucleic acid probes may be used for predicting, measuring and
XX   displaying gene expression in samples derived from human foetal liver. The
XX   present sequence is a single exon nucleic acid probe of the invention.
XX   Note: The sequence data for this patent did not form part of the printed
XX   specification, but was obtained in electronic format directly from WIPO
XX   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
XX
Query Match      3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      3185 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
Db      137 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
XX
QY      3245 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 3304
Db      77 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 18
XX
QY      3305 ATGGAGGGAATGTCTAG 3321
Db      17 ATGGAGGGAATGTCTAG 1
XX
RESULT 62
AAI48309/c
ID   AAI48309 standard; DNA; 137 BP.
XX
AC   AAI48309;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Probe #16995 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
XX   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
XX
PR   26-MAY-2000; 2000US-0207456P.
XX
PR   30-JUN-2000; 2000US-00608408.
XX
PR   03-AUG-2000; 2000US-00632366.
XX
PR   21-SEP-2000; 2000US-0234687P.
XX
PR   27-SEP-2000; 2000US-0236359P.
XX
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
XX   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 16995; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENPs).
XX   The present sequence is one such probe. The probes are useful for
XX   producing a microarray for predicting, measuring and displaying gene
XX   expression in samples derived from human placenta. The probes are useful
XX   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
XX
Query Match      3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      3185 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
Db      137 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
XX
QY      3245 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 3304
Db      77 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 18
XX
QY      3305 ATGGAGGGAATGTCTAG 3321
Db      17 ATGGAGGGAATGTCTAG 1
XX
RESULT 63
ABA50165/c
ID   ABA50165 standard; DNA; 137 BP.
XX
AC   ABA50165;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #8860.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;

```

cancer; ss.
Homo sapiens.
WO200157271-A2.
09-AUG-2001.
30-JAN-2001; 2001WO-US000662.
04-FEB-2000; 2000US-0180312P.
26-MAY-2000; 2000US-0207456P.
30-JUN-2000; 2000US-00608408.
03-AUG-2000; 2000US-00632366.
21-SEP-2000; 2000US-0234687P.
27-SEP-2000; 2000US-0236359P.
04-OCT-2000; 2000GB-00024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-496933/54.
New spatially-addressable set of single exon nucleic acid probes, useful for measuring gene expression in sample derived from human breast, comprises number of single exon nucleic acid probes.
Claim 4; SEQ ID NO 8860; 327pp + Sequence Listing; English.
The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human breast and BT 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for verifying the expression of regions of genomic DNA predicted to encode proteins. They are useful for gene discovery, and for determining predisposition and/or assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far less bias than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The present sequence is a single exon nucleic acid probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
Query Match 3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3185 CTTGCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGA 3244
Db 137 CTTGCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGA 78
QY 3245 ACCTGACCACAGGCAAGCACTGGAGACCTGCAATATCTGGCTTCTACGGTATCCACCA 3304
Db 77 ACCTGACCACAGGCAAGCACTGGAGACCTGCAATATCTGGCTTCTACGGTATCCACCA 18
QY 3305 ATGGAGGGAATGTCTAG 3321
Db 17 ATGGAGGGAATGTCTAG 1
RESULT 64
ABA35129/C
ID ABA35129 standard; DNA; 137 BP.
XX
AC ABA35129;

XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #13595 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PI WPI; 2001-488899/53.
XX
DR Single exon nucleic acid probes for analyzing gene expression in human hearts.
XX
PS Claim 4; SEQ ID NO 13595; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
Query Match 3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3185 CTTGCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGA 3244
Db 137 CTTGCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGA 78
QY 3245 ACCTGACCACAGGCAAGCACTGGAGACCTGCAATATCTGGCTTCTACGGTATCCACCA 3304
Db 77 ACCTGACCACAGGCAAGCACTGGAGACCTGCAATATCTGGCTTCTACGGTATCCACCA 18
QY 3305 ATGGAGGGAATGTCTAG 3321
Db 17 ATGGAGGGAATGTCTAG 1
RESULT 65
AAK42239/C
ID AAK42239 standard; DNA; 137 BP.
XX
AC AAK42239;
XX

DT XX 06-NOV-2001 (first entry)
 DE XX Human bone marrow expressed single exon probe SEQ ID NO: 16796.
 XX DE Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX OS Homo sapiens.
 XX PN WO200157276-A2.
 XX PD 09-AUG-2001.
 XX XX
 XX PF 30-JAN-2001; 2001WO-US000668.
 XX XX
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX XX WPI; 2001-488900/53.
 XX DR
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX PT gene expression in human bone marrow.
 XX XX
 XX PS Example 4; SEQ ID NO 16796; 658pp + Sequence Listing; English.
 XX XX
 XX CC The present invention provides a number of single exon nucleic acid
 XX CC probes which are derived from genomic sequences expressed in the human
 XX CC bone marrow. They can be used to measure gene expression in bone marrow
 XX CC samples, which may enable the improved diagnosis and treatment of cancers
 XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 XX CC the probes of the invention
 XX XX
 XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
 Query Match 3.2%; Score 137; DB 4; Length 137;
 Best Local Similarity 100.0%; Pred. No. 3.2e-53;
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3185 CTGCGCAATGCAACGGCCACAGTAAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
 Db 137 CTGCGCAATGCAACGGCCACAGTAAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
 QY 3245 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 3304
 Db 77 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 18
 QY 3305 ATGGAGGGAATGTCTAG 3321
 Db 17 ATGGAGGGAATGTCTAG 1
 RESULT 66
 AAK16486/c
 ID AAK16486 standard; DNA; 137 BP.
 XX AC AAK16486;
 XX XX
 XX DT 05-NOV-2001 (first entry)
 XX DE Human brain expressed single exon probe SEQ ID NO: 16477.
 XX DE Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
 KW ss.

XX OS Homo sapiens.
 XX PN WO200157275-A2.
 XX XX
 XX PD 09-AUG-2001.
 XX XX
 XX PF 30-JAN-2001; 2001WO-US000667.
 XX XX
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX XX WPI; 2001-483446/52.
 XX DR
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 XX PT brains.
 XX XX
 XX PS Example 4; SEQ ID NO 16477; 650pp + Sequence Listing; English.
 XX XX
 XX CC The present invention provides a number of single exon nucleic acid
 XX CC probes which are derived from genomic sequences expressed in the human
 XX CC brain. They can be used to measure gene expression in brain cell samples,
 XX CC which may enable the diagnosis and improved treatment of nervous system,
 XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 XX CC epilepsy and cancers. The present sequence is one of the probes of the
 XX CC invention
 XX XX
 XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
 Query Match 3.2%; Score 137; DB 4; Length 137;
 Best Local Similarity 100.0%; Pred. No. 3.2e-53;
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3185 CTGCGCAATGCAACGGCCACAGTAAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
 Db 137 CTGCGCAATGCAACGGCCACAGTAAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
 QY 3245 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 3304
 Db 77 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 18
 QY 3305 ATGGAGGGAATGTCTAG 3321
 Db 17 ATGGAGGGAATGTCTAG 1
 RESULT 67
 ABS41845/c
 ID ABS41845 standard; DNA; 137 BP.
 XX AC ABS41845;
 XX XX
 XX DT 25-FEB-2003 (first entry)
 XX DE Human liver single exon probe, SEQ ID No 16835.
 XX DE Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX OS Homo sapiens.
 XX PN WO200157273-A2.
 XX XX

PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000664.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX Claim 4; SEQ ID NO 16835; 658pp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. AB825011-AB851005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
Query Match 3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3185 CTGGCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
DB 137 CTGGCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
QY 3245 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCACCA 3304
DB 77 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCACCA 18
QY 3305 ATGGAGGGAATGTGAG 3321
DB 17 ATGGAGGGAATGTGAG 1
RESULT 68
AAI08671/c
ID AAI08671 standard; DNA; 137 BP.
XX AAI08671;
XX AAI08671;
XX 09-OCT-2001 (first entry)
XX Probe #8662 used to measure gene expression in human breast sample.
XX Probe; human; breast disease; breast cancer; development disorder; ss;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX Homo sapiens.
XX WO200157270-A2.
XX

XX 09-AUG-2001.
XX 29-JAN-2001; 2001WO-US000661.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX Novel single exon nucleic acid probe used to measuring gene expression in
XX a human breast.
XX Claim 25; SEQ ID NO 8662; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes.
XX The present sequence is one such probe. The probes are useful for
XX measuring human gene expression in a human breast sample, where the probe
XX hybridises at high stringency to a nucleic acid expressed in the human
XX breast. The probes are useful for predicting, diagnosing, grading,
XX staging, monitoring and prognosing diseases of the human breast.
XX Particularly those diseases with polygenic aetiology. The diseases
XX include: breast cancer, disorders of development, inflammatory diseases
XX of the breast, fibrocystic changes, proliferative breast disease and non-
XX carcinoma tumours. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
Query Match 3.2%; Score 137; DB 5; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3185 CTGGCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
DB 137 CTGGCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
QY 3245 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCACCA 3304
DB 77 ACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCACCA 18
QY 3305 ATGGAGGGAATGTGAG 3321
DB 17 ATGGAGGGAATGTGAG 1
RESULT 69
ABS16289/c
ID ABS16289 standard; DNA; 137 BP.
XX ABS16289;
XX ABS16289;
XX 19-AUG-2002 (first entry)
XX Human genome-derived single exon probe ORF from lung SEQ ID No 16280.
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX

KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.

PN -WO200186003-A2.

XX 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples.

XX Claim 4; SEQ ID NO 16280; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Judak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Query Match

Best Local Similarity 3.2%; Score 137; DB 6; Length 137;

Matches 137; Conservative 100.0%; Pred. No. 3.2e-53;

Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3185 CTTGCCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
 DB 137 CTTGCCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
 QY 3245 ACTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3304
 DB 77 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 18
 QY 3305 ATGGAGGGAATGTCTAG 3321
 DB 17 ATGGAGGGAATGTCTAG 1

RESULT 70

AAAL23548/c

ID AAL23548 standard; cDNA; 482 BP.

XX AAL23548;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 16005.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 2922; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
 CC (AA07544-AA026789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity

XX Sequence 482 BP; 106 A; 133 C; 126 G; 117 T; 0 U; 0 Other;

XX Query Match 3.0%; Score 130; DB 4; Length 482;

XX Best Local Similarity 100.0%; Pred. No. 6e-50;

XX Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2848 ATGTGGTGCAGCAACATGAGCAGTGTGTGACCTCCAAATGCTATGTGGCTCTTCCCT 2907

DB 130 ATGTGGTGCAGCAACATGAGCAGTGTGTGACCTCCAAATGCTATGTGGCTCTTCCCT 71

CC genes listed in the specification are useful as diagnostic or toxicity
 CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. These
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.

XX Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;

Query Match 1.4%; Score 62; DB 9; Length 8739;
 Best Local Similarity 100.0%; Pred. No. 3.7e-18;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 367 GCGTGGTGGGAGCAATGCCAGCACTCGGGGCGCGCTTCAGACTAAGTGGATCTTCT 426
 Db 436 GCGTGGTGGGAGCAATGCCAGCACTCGGGGCGCGCTTCAGACTAAGTGGATCTTCT 495
 QY 427 GG 428
 Db 496 GG 497

RESULT 74
 ABN42673
 ID ABN42673 standard; DNA; 60 BP.
 AC ABN42673;
 XX
 DT 15-JUL-2002 (first entry)
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:15421.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX Homo sapiens.
 OS
 PN WC200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-IB001903.
 XX
 PR 28-JUL-2000; 2000US-0221607P.
 PR 02-MAY-2001; 2001US-0287724P.
 XX
 PA (COMP-) COMPUGEN INC.
 XX
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX
 DR WPI; 2002-257383/30.
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a
 PT genome, useful for detecting tissue-, pathology-, and developmental-
 PT specific genes.
 XX
 PS Example 1; SEQ ID NO 15421; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
 CC)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridising selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a particular

CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC developmental specific genes; and to detect RNA transcripts and splice
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 60 BP; 19 A; 14 C; 13 G; 14 T; 0 U; 0 Other;
 Query Match 1.4%; Score 60; DB 6; Length 60;
 Best Local Similarity 100.0%; Pred. No. 3.3e-17;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3547 GTGGCTACTCTCGACGAACAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAT 3606
 Db 1 GTGGCTACTCTCGACGAACAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAT 60

RESULT 74
 AAZ91925
 ID AAZ91925 standard; DNA; 90050 BP.
 AC AAZ91925;
 XX
 DT 08-JUN-2000 (first entry)
 DE Wild type (C57BL/6J) murine mahogany protein genomic sequence c110/111.
 XX Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
 KW weight regulation; cell therapy; body weight disorder; cachexia;
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;
 KW hyperphagia; Anticobesity; antianorexic; anticachexic; ds.
 XX Mus sp.

XX WC2000005373-A2.
 XX
 PD 03-FEB-2000.
 XX
 PF 21-JUL-1999; 99WO-US016484.
 XX
 PR 21-JUL-1998; 98US-0093630P.
 PR 20-OCT-1998; 98US-0104978P.
 PR 05-FEB-1999; 99US-00245041.
 XX
 PA (MILL-) MILLENIUM PHARM INC.
 XX
 PI Moore K, Nagle DL;
 XX
 DR WPI; 2000-195103/17.

PT New human and murine mahogany genes, useful, e.g. for diagnosis and
 PT treatment of body weight disorders.

XX Example; Fig 3d; 188pp; English.

XX This sequence represents a murine mahogany gene of the invention. The
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
 CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
 CC therapeutics; (iii) as a source of diagnostic probes and primers for
 CC detecting expression of mg genes or mutations, regulatory defects, in
 CC this gene, or for isolation of related sequences; and (iv) in (cell-
 CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
 CC identify other (extra)cellular products involved in weight regulation,
 CC and to screen for agents that disrupt interaction between (II) and other
 CC macromolecules. The Ab are used to detect abnormal levels (or function)
 CC of (II) for diagnosis, prognosis or monitoring of treatment; to
 CC evaluate (II)-expressing cells intended for cell therapy, and as
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the

CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 90050 BP; 25307 A; 18193 C; 19454 G; 27096 T; 0 U; 0 Other;
Query Match 1.3%; Score 55; DB 3; Length 90050;
Best Local Similarity 100.0%; Pred. No. 6.9e-15;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3130 TCCAGCATGCTCTAGAGGACGACGATACACTGCTCTTCATTCACTGTCACG 3184
DB 31840 TCCAGCATGCTCTAGAGGACGACGATACAACTGGTCTTTCATTCACTGTCACG 31894
RESULT 75
AAZ91918
ID AAZ91918 standard; cDNA; 1051 BP.
XX
AC AAZ91918;
XX
DT 08-JUN-2000 (first entry)
XX
DE Murine mahogany protein coding sequence akml004.
XX
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Mus sp.
XX
FN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 95WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore X, Nagle DL;
XX
DR WPI; 2000-195103/17.
XX
DR P-PSDB; AAY81805.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 9a; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (ii) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (iii)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,

CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 1051 BP; 306 A; 243 C; 313 G; 189 T; 0 U; 0 Other;
Query Match 1.2%; Score 53; DB 3; Length 1051;
Best Local Similarity 100.0%; Pred. No. 6.1e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 367 GGCTGGTGGCGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGG 419
DB 571 GGCTGGTGGCGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGG 623

Search completed: March 1, 2004, 13:39:57
Job time : 1480 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2004, 13:27:38 ; Search time 1409 Seconds

(without alignments)

10983.699 Million cell updates/sec

Title: US-09-787-097-13

Perfect score: 4290

Sequence: 1 atgtgtccgcagcggcgc.....agctggacctgcattga 4290

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2353733 seqs, 180373377 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3774	88.0	5589	US-09-893-238-14	Sequence 14, Appl
2	3285	76.6	4072	US-09-893-238-16	Sequence 16, Appl
3	1837	42.8	2625	US-09-893-238-18	Sequence 18, Appl
4	1510	35.2	6370	US-09-893-238-12	Sequence 12, Appl
5	291	6.8	8827	US-09-893-238-1	Sequence 1, Appl
6	246	5.7	246	US-09-736-457-966	Sequence 966, App
7	246	5.7	246	US-09-902-941-966	Sequence 966, App
8	246	5.7	246	US-09-849-626-966	Sequence 966, App
9	246	5.7	246	US-10-017-754-966	Sequence 966, App
10	246	5.7	246	US-10-113-872-966	Sequence 966, App
11	234	5.5	207433	US-10-277-216-5	Sequence 5, Appl
12	234	5.5	207433	US-10-126-022-5	Sequence 5, Appl
13	228	5.3	625	US-10-198-846-5869	Sequence 5869, App
14	210	4.9	531	US-09-864-761-15900	Sequence 15900, A
15	193	4.5	193	US-09-864-761-32404	Sequence 32404, A

182	4.2	182	9	US-09-864-761-21460	Sequence 21460, A
173	4.0	375	9	US-09-864-761-4720	Sequence 4720, Ap
141	3.3	402	9	US-09-864-761-3681	Sequence 3681, Ap
139	3.2	212	9	US-09-796-692-7610	Sequence 7610, Ap
139	3.2	212	14	US-10-040-862-7610	Sequence 7610, Ap
139	3.2	212	15	US-10-057-475B-7610	Sequence 7610, Ap
139	3.2	212	15	US-10-154-884B-7610	Sequence 7610, Ap
137	3.2	137	9	US-09-864-761-20449	Sequence 20449, A
86	2.0	594	15	US-10-027-632-264639	Sequence 264639, A
86	2.0	594	15	US-10-027-632-264640	Sequence 264640, A
79	1.8	629	15	US-10-027-632-270240	Sequence 270240, A
60	1.4	60	10	US-09-908-975-15421	Sequence 15421, A
55	1.3	90050	9	US-09-893-238-5	Sequence 5, Appl
53	1.2	1051	9	US-09-893-238-10	Sequence 10, Appl
47	1.1	2419	9	US-09-893-238-8	Sequence 8, Appl
38	0.9	468	15	US-10-027-632-108988	Sequence 108988, A
35	0.8	5973	9	US-09-893-238-4	Sequence 4, Appl
32	0.7	65	10	US-09-908-975-29990	Sequence 29990, A
32	0.7	17056	9	US-09-893-238-3	Sequence 3, Appl
29	0.7	29	9	US-09-893-238-23	Sequence 23, Appl
29	0.7	1332	15	US-10-419-296-1	Sequence 1, Appl
29	0.7	1622	14	US-10-029-386-22666	Sequence 22666, A
27	0.6	5312	14	US-10-311-455-169	Sequence 169, App
24	0.6	24	9	US-09-893-238-48	Sequence 48, Appl
24	0.6	498	9	US-09-783-590-7412	Sequence 7412, Ap
24	0.6	784	12	US-10-424-599-106383	Sequence 106383, A
24	0.6	899	12	US-10-425-114-13318	Sequence 13318, A
24	0.6	1003	12	US-10-425-114-29183	Sequence 29183, A
24	0.6	1097	12	US-10-424-599-10899	Sequence 10899, A
24	0.6	1372	12	US-10-424-599-547	Sequence 547, App
24	0.6	1756	12	US-10-424-599-21459	Sequence 21459, A
24	0.6	2180	10	US-09-814-353-21391	Sequence 21391, A
24	0.6	2294	15	US-10-437-171-3	Sequence 3, Appl
24	0.6	3334	15	US-10-437-171-1	Sequence 1, Appl
24	0.6	4964	12	US-10-424-599-56004	Sequence 56004, A
24	0.6	5085	14	US-10-198-846-9854	Sequence 9854, Ap
24	0.6	7386	10	US-09-819-104A-6	Sequence 6, Appl
24	0.6	7521	10	US-09-819-104A-3	Sequence 3, Appl
24	0.6	7554	15	US-10-414-692-14	Sequence 14, Appl
24	0.6	8544	15	US-09-819-104A-4	Sequence 4, Appl
24	0.6	8561	15	US-10-174-014-11	Sequence 11, Appl
24	0.6	8686	15	US-09-819-104A-1	Sequence 1, Appl
24	0.6	8686	15	US-10-174-014-4	Sequence 4, Appl
24	0.6	14859	15	US-10-051-874-55	Sequence 55, Appl
24	0.6	15845	15	US-10-085-198-111	Sequence 11, Appl
24	0.6	221000	15	US-10-174-014-12	Sequence 12, Appl
23	0.5	29	9	US-09-893-238-24	Sequence 24, Appl
23	0.5	104	15	US-10-260-238-1970	Sequence 1970, Ap
23	0.5	221	9	US-09-893-238-7	Sequence 7, Appl
23	0.5	226	9	US-09-893-238-6	Sequence 6, Appl
23	0.5	500	14	US-10-029-386-11303	Sequence 11303, A
23	0.5	510	15	US-10-260-238-5694	Sequence 5694, Ap
23	0.5	625	9	US-09-864-761-19977	Sequence 19977, A
23	0.5	728	12	US-10-424-599-110798	Sequence 110798, A
23	0.5	900	12	US-10-425-114-20706	Sequence 20706, A
23	0.5	1320	15	US-10-260-238-2125	Sequence 2125, Ap
23	0.5	1379	9	US-09-864-761-3198	Sequence 3198, Ap
23	0.5	2000	15	US-10-260-238-1864	Sequence 1864, Ap
23	0.5	2200	15	US-10-260-238-1920	Sequence 1920, Ap
23	0.5	2264	15	US-10-108-260A-790	Sequence 790, App
23	0.5	6733	16	US-10-197-824-1	Sequence 1, Appl
23	0.5	300000	14	US-10-262-552-33	Sequence 33, Appl
22	0.5	26	9	US-09-893-238-30	Sequence 30, Appl
22	0.5	188	10	US-09-764-881-177	Sequence 177, App
22	0.5	238	15	US-10-242-747-177	Sequence 177, App
22	0.5	238	9	US-09-864-761-24079	Sequence 24079, A
22	0.5	293	9	US-09-864-761-18923	Sequence 18923, A
22	0.5	396	14	US-10-029-386-17961	Sequence 17961, A
22	0.5	418	12	US-10-424-599-104491	Sequence 104491, A
22	0.5	435	15	US-10-027-632-181730	Sequence 181730, A
22	0.5	435	15	US-10-027-632-308272	Sequence 308272, A
22	0.5	436	9	US-09-728-445-698	Sequence 698, App
22	0.5	437	12	US-10-424-599-107947	Sequence 107947, A

c 89 22 0.5 459 9 US-09-864-761-2182 Sequence 2182, Ap
90 22 0.5 459 9 US-09-981-876-36 Sequence 36, Appl
91 22 0.5 459 10 US-09-148-545-36 Sequence 36, Appl
92 22 0.5 480 10 US-09-770-961-950 Sequence 950, App
93 22 0.5 554 9 US-09-864-761-7357 Sequence 7357, Ap
c 94 22 0.5 585 14 US-10-029-386-4261 Sequence 4261, Ap
c 95 22 0.5 607 12 US-10-029-114-22348 Sequence 22348, A
c 96 22 0.5 616 14 US-10-425-114-22348 Sequence 22348, A
c 97 22 0.5 616 15 US-10-242-535A-1932 Sequence 1932, A
c 98 22 0.5 654 14 US-10-156-761-2741 Sequence 2741, Ap
c 99 22 0.5 665 12 US-10-424-599-97537 Sequence 97537, A
c 100 22 0.5 775 12 US-10-424-599-16675 Sequence 16675, A

ALIGNMENTS

RESULT 1
US-09-893-238-14
; Sequence 14, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893, 238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245, 041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093, 630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104, 978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-14

Query Match 88.0%; Score 3774; DB 9; Length 8589;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3974; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 313 TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCCAGTCGCTGCCCCCGCGGTGG 372
DB 310 TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCCAGTCGCTGCCCCCGCGGTGG 369
QY 373 GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAATGATCTTCGGGTTT 432
DB 370 GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAATGATCTTCGGGTTT 429
QY 433 GTGACAGATGACCTGGGAAATATATAAACAAGAGTCAGTCGCTCATTTGAAGGA 492
DB 430 GTGACAGATGACCTGGGAAATATATAAACAAGAGTCAGTCGCTCATTTGAAGGA 489
QY 493 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTGTAAGAGTGTAGTGGAC 552
DB 490 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTGTAAGAGTGTAGTGGAC 549
QY 553 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC 612
DB 550 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC 609
QY 613 CTCATTGTTTCTGAGAGAGATGGCAATGAGACTGTTCCTCGAGGTGTTGCCACATCAGGT 672
DB 610 CTCATTGTTTCTGAGAGAGATGGCAATGAGACTGTTCCTCGAGGTGTTGCCACATCAGGT 669
QY 673 TATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTTACCTGGATTTAATATTACT 732

DB 670 TATGCTTGTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGAATTAATATTACT 729
QY 733 TACAGTTTGTGATGTGTCCTCAATTAACCTGCTCAGCCGAGGAGAGTGAAGATCAGTAAT 792
DB 730 TACAGTTTGTGATGTGTCCTCAATTAACCTGCTCAGCCGAGGAGAGTGAAGATCAGTAAT 789
QY 793 AGCAGCGAAACTGTTCAATGTGAATGTTTCTGAAAACTGGAAAGGTGAAGCATGTGACATT 852
DB 790 AGCAGCGAATGTTGTTGAATGTGAATGTTTCTGAAAACTGGAAAGGTGAAGCATGTGACATT 849
QY 853 CCTCACTGTACAGACAACCTGCTGCTCTCATCGAGCATCTGCAATTCAGTCAATGTC 912
DB 850 CCTCACTGTACAGACAACCTGCTGCTCTCATCGAGCATCTGCAATTCAGTCAATGTC 909
QY 913 AGAGATGCTCTCTGCTTCTCAGACTGGCAGGCTCTCGATGTTTCACTTCCTGTACCACT 972
DB 910 AGAGATGCTCTCTGCTTCTCAGACTGGCAGGCTCTCGATGTTTCACTTCCTGTACCACT 969
QY 973 AACAGTCAATTTGGACTCGAGAGAAATTTCTAATTAAGCTCCCGAGCATCTCAT 1032
DB 970 AACAGTCAATTTGGACTCGAGAGAAATTTCTAATTAAGCTCCCGAGCATCTCAT 1029
QY 1033 AAAGCTGTGCTCAATGGAAACATTTATGCGGTGTTTGGAGGATATATGTTCAACCACTCA 1092
DB 1030 AAAGCTGTGCTCAATGGAAACATTTATGCGGTGTTTGGAGGATATATGTTCAACCACTCA 1089
QY 1093 GATTATAACATGTTTCTAGCGTATGACCTTGTCTAGGAGTGGCTTCCACTAAACCGT 1152
DB 1090 GATTATAACATGTTTCTAGCGTATGACCTTGTCTAGGAGTGGCTTCCACTAAACCGT 1149
QY 1153 TCTGTGAACAATGCTGTTTGTAGATATGCTGCTCTTTGGCAATATACAAGATATAAT 1212
DB 1150 TCTGTGAACAATGCTGTTTGTAGATATGCTGCTCTTTGGCAATATACAAGATATAAT 1209
QY 1213 TACATGATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAGTGTGAGTTT 1272
DB 1210 TACATGATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAGTGTGAGTTT 1269
QY 1273 CACATTCATTAAGTATGCTGGTGTGTTGACCCCTAAGGCAAGGACGATGTCAGTG 1332
DB 1270 CACATTCATTAAGTATGCTGGTGTGTTGACCCCTAAGGCAAGGACGATGTCAGTG 1329
QY 1333 GTTGGCACTCTGCACACATTTCTTACACTGAAGAATGCGCAGTGTGTCATGTCGTATC 1392
DB 1330 GTTGGCACTCTGCACACATTTCTTACACTGAAGAATGCGCAGTGTGTCATGTCGTATC 1389
QY 1393 TTTGTCATGCTGCTCTCTATGAGATATATAGCAATGTGACGAAATATGATTTGGATAG 1452
DB 1390 TTTGTCATGCTGCTCTCTATGAGATATATAGCAATGTGACGAAATATGATTTGGATAG 1449
QY 1453 AACACATGAGTATATACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC 1512
DB 1450 AACACATGAGTATATACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC 1509
QY 1513 AGTGTTTACGACCAATAGGACACAGGCGCTTATAGTTCATGTTGGCTACAAGCTTTTCA 1572
DB 1510 AGTGTTTACGACCAATAGGACACAGGCGCTTATAGTTCATGTTGGCTACAAGCTTTTCA 1569
QY 1573 GCGMATAGTACCGGTTGCGAGATGATCTTACCGATATGATGTGATACCCAGATGTGG 1632
DB 1570 GCGMATAGTACCGGTTGCGAGATGATCTTACCGATATGATGTGATACCCAGATGTGG 1629
QY 1633 ACCATTTCTTAAGGACACAGCGGATTTTCCGTTCATTTGTCACACAGCTGTGATGTGGA 1692
DB 1630 ACCATTTCTTAAGGACACAGCGGATTTTCCGTTCATTTGTCACACAGCTGTGATGTGGA 1689
QY 1693 ACCATGCTGTTTGGGGGAAACACACACATGACATCTCTAGCCATGCGCCCAAA 1752
DB 1690 ACCATGCTGTTTGGGGGAAACACACACATGACATCTCTAGCCATGCGCCCAAA 1749
QY 1753 TGTCTCTTTCAGATTTTATGCGCTTATGACATTTGCTGACCGCTGTGCTGCTGCTTCCC 1812
DB 1750 TGTCTCTTTCAGATTTTATGCGCTTATGACATTTGCTGACCGCTGTGCTGCTGCTTCCC 1809

Db	3970	AGCGTCCCTTGGCTCTGTAAATGTCGCTTGGAAACAGATGAGGAGCCTCCTGATCTT	4029	
Qy	4033	ATTGGGGGAGTATAAGAGCTGTTCCCAACCCATTGCACTGGAGCGGTGTTTGGCAAC	4092	
Db	4030	ATTGGGGGAGTATAAGAGCTGTTCCCAACCCATTGCACTGGAGCGGTGTTTGGCAAC	4089	
Qy	4093	AAAGCGCTGTCCCTCTCTGCTGTTGTGAGGCTCCCTCGAGCGCTGGGTGGCATCCCTCCT	4152	
Db	4090	AAAGCGCTGTCCCTCTCTGCTGTTGTGAGGCTCCCTCGAGCGCTGGGTGGCATCCCTCCT	4149	
Qy	4153	CCTGGGAGTCAAGTCTGCTGTTGGGCGAGCGCCCTGGTGGACATTTCTCAGCAGATGCCG	4212	
Db	4150	CCTGGGAGTCAAGTCTGCTGTTGGGCGAGCGCCCTGGTGGACATTTCTCAGCAGATGCCG	4209	
Qy	4213	ATAGTGTACAAGGAGAAGTCAGGAGCGGTGAGAAACCGGAGCAGAGCGCCCTGCACAG	4272	
Db	4210	ATAGTGTACAAGGAGAAGTCAGGAGCGGTGAGAAACCGGAGCAGAGCGCCCTGCACAG	4269	
Qy	4273	CCTGGGACCTGCATCTGA	4290	
Db	4270	CCTGGGACCTGCATCTGA	4287	
RESULT 2				
US-09-893-238-16				
; Sequence 16, Application US/09893238				
; Patent No. US20020150973A1				
; GENERAL INFORMATION:				
; APPLICANT: Moore, K.				
; APPLICANT: Nagle, D.				
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND				
; FILE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY				
; CURRENT APPLICATION NUMBER: US/09/893,238				
; CURRENT FILING DATE: 2001-06-27				
; PRIOR APPLICATION NUMBER: 09/245,041				
; PRIOR FILING DATE: 1999-02-05				
; PRIOR APPLICATION NUMBER: 60/093,630				
; PRIOR FILING DATE: 1998-07-21				
; PRIOR APPLICATION NUMBER: 60/104,978				
; PRIOR FILING DATE: 1998-10-20				
; NUMBER OF SEQ ID NOS: 129				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 16				
; LENGTH: 4072				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-09-893-238-16				
Query Match 76.6%; Score 3285; DB 9; Length 4072;				
Best Local Similarity 99.9%; Pred. No. 0;				
Matches 3485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				
Qy	313	TGTTCAACGGCGGTGCGTCAACCCCTGGCAACCGGCGAGTGGCTGCTCCCGCGCGGTGG	372	
Db	310	TGTTCAACGGCGGTGCGTCAACCCCTGGCAACCGGCGAGTGGCTGCTCCCGCGCGGTGG	369	
Qy	373	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTTAAGTCTCTGGGTTT	432	
Db	370	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTTAAGTCTCTGGGTTT	429	
Qy	433	GTGACAGATGACCTGGAAATATAAATACAAACGAAGTGCACGTGGCTCATTTGAAGGA	492	
Db	430	GTGACAGATGACCTGGAAATATAAATACAAACGAAGTGCACGTGGCTCATTTGAAGGA	489	
Qy	493	CAGCCAAATAGAAATAGACTTCGTTTCAATCTTTCGTACAGAGTGTAGTGGAC	552	
Db	490	CAGCCAAATAGAAATAGACTTCGTTTCAATCTTTCGTACAGAGTGTAGTGGAC	549	
Qy	553	CATTATATGTTATGATGGGAGTCAATTTATGCACCGGTAGTGTGCTGCTATTAGTGGC	612	
Db	550	CATTATATGTTATGATGGGAGTCAATTTATGCACCGGTAGTGTGCTGCTATTAGTGGC	609	

Qy	613	CTCATTGTTCTTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT	672	
Db	610	CTCATTGTTCTTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT	669	
Qy	673	TATGCTTCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATTAATATTACT	732	
Db	670	TATGCTTCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATTAATATTACT	729	
Qy	733	TACAGTTTGTATGATGCTGCTCAAAATAACTGCTCAGGCCGAGGAGAGTGTAAAGTCAAGTAAT	792	
Db	730	TACAGTTTGTATGATGCTGCTCAAAATAACTGCTCAGGCCGAGGAGAGTGTAAAGTCAAGTAAT	789	
Qy	793	ACGACGCAAACTGTTGAATGTAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACATTT	852	
Db	790	ACGACGCAAACTGTTGAATGTAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACATTT	849	
Qy	853	CCTCACTGTACAGACAACTGTGTTTCTCATCGAGGCATCTGCAATTCAGTTCCTGTACAGCT	912	
Db	850	CCTCACTGTACAGACAACTGTGTTTCTCATCGAGGCATCTGCAATTCAGTTCCTGTACAGCT	909	
Qy	913	AGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGATGTTTCTGTTCCTGTACAGCT	972	
Db	910	AGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGATGTTTCTGTTCCTGTACAGCT	969	
Qy	973	AACCACTGTTTGGACTCGAGAGCAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1032	
Db	970	AACCACTGTTTGGACTCGAGAGCAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1029	
Qy	1033	AAAGCTGTGTTCAATGGAACCAATTTATGTTGGGTGTTTGGAGGATATATGTTCAACCACTCA	1092	
Db	1030	AAAGCTGTGTTCAATGGAACCAATTTATGTTGGGTGTTTGGAGGATATATGTTCAACCACTCA	1089	
Qy	1093	GATTATAACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGT	1152	
Db	1090	GATTATAACATGTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGT	1149	
Qy	1153	TCTGTGAACAATGTTGTTTGTAGATATGTTGATGTTTCTTTGGCATATATCAAGATATAAAT	1212	
Db	1150	TCTGTGAACAATGTTGTTTGTAGATATGTTGATGTTTCTTTGGCATATATCAAGATATAAAT	1209	
Qy	1213	TACATGTTATGAGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTCAGAGTTTT	1272	
Db	1210	TACATGTTATGAGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTCAGAGTTTT	1269	
Qy	1273	CACATTCATATGAGTTCATGCTGTTGTTGACCTTAAAGCAAGGAGCAGTATGCAAGT	1332	
Db	1270	CACATTCATATGAGTTCATGCTGTTGTTGACCTTAAAGCAAGGAGCAGTATGCAAGT	1329	
Qy	1333	GTGGGCACTCTGCACACATTTTACACTGAAGATGCGCGAGTGGTTCATGCTGGTCAATC	1392	
Db	1330	GTGGGCACTCTGCACACATTTTACACTGAAGATGCGCGAGTGGTTCATGCTGGTCAATC	1389	
Qy	1393	TTTGTGCTACCTGCTCTCTATGATATATAGCAATGTGAGGATATGATGTTGGATTAAG	1452	
Db	1390	TTTGTGCTACCTGCTCTCTATGATATATAGCAATGTGAGGATATGATGTTGGATTAAG	1449	
Qy	1453	AACACATGAGATATATTACACACCGGAGTGGCTTGTGCAAGGGGGTTACGGCCATAGC	1512	
Db	1450	AACACATGAGATATATTACACACCGGAGTGGCTTGTGCAAGGGGGTTACGGCCATAGC	1509	
Qy	1513	AGTGTGTTACGACATPAGGACCGAGGCCCTTACGTTTCAATGTTGGCTCAAGGCTTTTCA	1572	
Db	1510	AGTGTGTTACGACATPAGGACCGAGGCCCTTACGTTTCAATGTTGGCTCAAGGCTTTTCA	1569	
Qy	1573	GCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTTGGATACCGAGATGTTG	1632	
Db	1570	GCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTTGGATACCGAGATGTTG	1629	
Qy	1633	ACCATTCTTAAGGACAGCGGATTTTCCGTTTCTGTTGACACAGCTGTGATGAGTGGGA	1692	
Db	1630	ACCATTCTTAAGGACAGCGGATTTTCCGTTTCTGTTGACACAGCTGTGATGAGTGGGA	1689	

US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-05-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Query Match 42.8%; Score 1837; DB 9; Length 2625;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2037; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	313	TGTGTCACAGCGGTCGCTGCAACCCCTGGCACCAGCGGAGTCGCTGCGCCCGCGCGCTGG	372
DB	310	TGTGTCACAGCGGTCGCTGCAACCCCTGGCACCAGCGGAGTCGCTGCGCCCGCGCGCTGG	369
QY	373	GTGGCGAGCAATCCAGCACTCGGGGGCGCGCTTCAGACTAACTGATCTTTGGGTTT	432
DB	370	GTGGCGAGCAATCCAGCACTCGGGGGCGCGCTTCAGACTAACTGATCTTTGGGTTT	429
QY	433	GTGACAGATGACCTGGGAATTAATAACAAAGAGTGCAGTGGCTCATTTAGGA	492
DB	430	GTGACAGATGACCTGGGAATTAATAACAAAGAGTGCAGTGGCTCATTTAGGA	489
QY	493	CAGCCAAATAGAAATACAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGAC	552
DB	490	CAGCCAAATAGAAATACAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGAC	549
QY	553	CATTATATGTTTATGATGGGACTCAATTTATGCAACGGCTAGTTGCTGATTTAGTGGC	612
DB	550	CATTATATGTTTATGATGGGACTCAATTTATGCAACGGCTAGTTGCTGATTTAGTGGC	609
QY	613	CTCATTGTTCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGGT	672
DB	610	CTCATTGTTCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGGT	669
QY	673	TATGCTTGCTGCATTTTTTTTATGATGCTGCTTATTAATTTGACTGATTTAATATCT	732
DB	670	TATGCTTGCTGCATTTTTTTTATGATGCTGCTTATTAATTTGACTGATTTAATATCT	729
QY	733	TACAGTTTTGATATGTCGCAAAATGCTCTCAGCGCAGAGAGTGTGAAGATCAGTAA	792
DB	730	TACAGTTTTGATATGTCGCAAAATGCTCTCAGCGCAGAGAGTGTGAAGATCAGTAA	789
QY	793	AGCAGCGAACTGTTGAATGTAATGTTCTGAAATCTGAAAGCTGAAGCATGTGACATT	852
DB	790	AGCAGCGAATCTGTTGAATGTAATGTTCTGAAATCTGAAAGCTGAAGCATGTGACATT	849
QY	853	CCTCACTGTACAGCAACTGTTGTTTCTTCATCGAGCATCTGCAATTCAGTGTATTC	912
DB	850	CCTCACTGTACAGCAACTGTTGTTTCTTCATCGAGCATCTGCAATTCAGTGTATTC	909
QY	913	AGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTTCCTGTACCACT	972
DB	910	AGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTTCCTGTACCACT	969

QY	973	AAACAGTCTATTTGGACTCGAGAGGAATATTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1032
DB	970	AAACAGTCTATTTGGACTCGAGAGGAATATTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1029
QY	1033	AAAGCTGTGTCATATGGAAACATTATGTGGGTGTGTGGAGGATATATCTTCAACACTCA	1092
DB	1030	AAAGCTGTGTCATATGGAAACATTATGTGGGTGTGTGGAGGATATATCTTCAACACTCA	1089
QY	1093	GATTATAACATGCTTCTAGCGTATGACCTTGTCTTCTAGGGAGTGGCTTCCACTAAACCGT	1152
DB	1090	GATTATAACATGCTTCTAGCGTATGACCTTGTCTTCTAGGGAGTGGCTTCCACTAAACCGT	1149
QY	1153	TCTGTGAACAATGTGGTGTGTAGATATGCTTCTTCTTGGCAATATATCAAGGATAAATTT	1212
DB	1150	TCTGTGAACAATGTGGTGTGTAGATATGCTTCTTCTTGGCAATATATCAAGGATAAATTT	1209
QY	1213	TACATGTATGAGGAGAAATTCGATCCAACTGGGAATGTGACCAATGAGTTCAGAGTTTTT	1272
DB	1210	TACATGTATGAGGAGAAATTCGATCCAACTGGGAATGTGACCAATGAGTTCAGAGTTTTT	1269
QY	1273	CACATTCATATAGTTCATGCGTGTGTGACCCCTTAAGGCAAGAGAGCATGTGACGTG	1332
DB	1270	CACATTCATATAGTTCATGCGTGTGTGACCCCTTAAGGCAAGAGAGCATGTGACGTG	1329
QY	1333	GTGGGCACTGTCACACATTTTACACTGAAGATGCGCGAGTGTCTATCTGTCTATC	1392
DB	1330	GTGGGCACTGTCACACATTTTACACTGAAGATGCGCGAGTGTCTATCTGTCTATC	1389
QY	1393	TTTGTGCTCACTGCGCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG	1452
DB	1390	TTTGTGCTCACTGCGCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG	1449
QY	1453	AAACATGAGGATATATACACACCAGGCTGCGCTTGTGCAAGGGGTTCAGGCCATAGC	1512
DB	1450	AAACATGAGGATATATACACACCAGGCTGCGCTTGTGCAAGGGGTTCAGGCCATAGC	1509
QY	1513	AGTGTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTTGGCTACAAGGCTTTTCACT	1572
DB	1510	AGTGTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTTGGCTACAAGGCTTTTCACT	1569
QY	1573	GCCAAATAGTACCGCTTTCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGG	1632
DB	1570	GCCAAATAGTACCGCTTTCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGG	1629
QY	1633	ACCATTCCTTAAGGACAGCGGATTTTTCGTTTACTTTGACACAGCTGTGATAGTGTGGA	1692
DB	1630	ACCATTCCTTAAGGACAGCGGATTTTTCGTTTACTTTGACACAGCTGTGATAGTGTGGA	1689
QY	1693	ACCATGCTGTTGTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGGCGCCAAA	1752
DB	1690	ACCATGCTGTTGTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGGCGCCAAA	1749
QY	1753	TGCTTCTCTTCAGATTTTATGCGCTTATGACATTTGCGCTGTCAGCGCTGCTGCTTCCC	1812
DB	1750	TGCTTCTCTTCAGATTTTATGCGCTTATGACATTTGCGCTGTCAGCGCTGCTGCTTCCC	1809
QY	1813	AGACTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAACAGC	1872
DB	1810	AGACTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAACAGC	1869
QY	1873	ACCATGTATGTTGCTGGTGTGTTTCAATAGTCTCTCTCCTCAGGACATCTCTGTTATCAC	1932
DB	1870	ACCATGTATGTTGCTGGTGTGTTTCAATAGTCTCTCTCCTCAGGACATCTCTGTTATCAC	1929
QY	1933	TCGGAAACAGTGTGATGCGCATCGGAGTCAAGCGGCTGTTTGTAGCAGCAGGACCTGGTATT	1992
DB	1930	TCGGAAACAGTGTGATGCGCATCGGAGTCAAGCGGCTGTTTGTAGCAGCAGGACCTGGTATT	1989
QY	1993	CGGTGTGTGTGAAACACAGGCTGCTTCAGTGTATCTGCTGGGCGCTGGCACTGATGAA	2052
DB	1990	CGGTGTGTGTGAAACACAGGCTGCTTCAGTGTATCTGCTGGGCGCTGGCACTGATGAA	2049
QY	2053	CAAGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTTGACCATGACAGATGT	2112

Db	2050	CAAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGT	QY	3081	GCCTTCGCAAGCCCTCAGAGAAATTTCTATCCAGAGCCCTGTCTCAATTCACAGCATGTG	3140
QY	2113	GACGAGCACACAGATTGTTACAGGTGTACAGGTGTACAGCCCAACCACTGCTGCTGCTCAAT	Db	848	GCCTTCGCAAGCCCTCAGAGAAATTTCTATCCAGAGCCCTGTCTCAATTCACAGCATGTG	907
Db	2110	GACGAGCACACAGATTGTTACAGGTGTACAGGTGTACAGCCCAACCACTGCTGCTGCTCAAT	QY	3141	TCCTAGAGGACAGACATACAACTGCTCTTTCTATTCCTGCTCCAGCTTCCCAATGCAACGG	3200
QY	2173	GACCATTTGTGTCCCAAGAACCAACAGCTGTCTCAGAGCCCGAGATCTCCATTTTAGGTAT	Db	908	TCCTAGAGGACAGACATACAACTGCTCTTTCTATTCCTGCTCCAGCTTCCCAATGCAACGG	967
Db	2170	GACCATTTGTGTCCCAAGAACCAACAGCTGTCTCAGAGCCCGAGATCTCCATTTTAGGTAT	QY	3201	CCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAAGGCAA	3260
QY	2233	GAGAAATGCCCAAGGATACCCCATGTACTACTGTAAACAGAGACCACTGACGAGGC	Db	968	CCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAAGGCAA	1027
Db	2230	GAGAAATGCCCAAGGATACCCCATGTACTACTGTAAACAGAGACCACTGACGAGGC	QY	3261	GCACTGCGAGACCTGCAATCTGGCTTCTPACGGTGTATCCCAATGAGGGGAAATGTCA	3320
QY	2293	TGTGCTTGGACCAAGAACTGCACTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCC	Db	1028	GCACTGCGAGACCTGCAATCTGGCTTCTPACGGTGTATCCCAATGAGGGGAAATGTCA	1087
Db	2290	TGTGCTTGGACCAAGAACTGCACTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCC	QY	3321	GCACTGCAAGTGCATTTGGACACGCTCTCTGTGCAACCAACACGCGCAAGTGTCTTG	3380
QY	2353	G 2353	Db	1088	GCACTGCAAGTGCATTTGGACACGCTCTCTGTGCAACCAACACGCGCAAGTGTCTTG	1147
Db	2350	G 2350	QY	3381	CACCAAGAGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCA	3440
RESULT 4						
US-09-893-238-12						
; Sequence 12, Application US/09893238						
; Patent No. US20020150973A1						
; GENERAL INFORMATION:						
; APPLICANT: Moore, K.						
; APPLICANT: Nagle, D.						
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND						
; FILE REFERENCE: 7853-237						
; CURRENT APPLICATION NUMBER: US/09/893,238						
; CURRENT FILING DATE: 2001-06-27						
; PRIOR APPLICATION NUMBER: 09/245,041						
; PRIOR FILING DATE: 1999-02-05						
; PRIOR FILING DATE: 1998-07-21						
; PRIOR FILING DATE: 1998-10-20						
; NUMBER OF SEQ ID NOS: 129						
; SOFTWARE: FastSeq for Windows Version 3.0						
; SEQ ID NO 12						
; LENGTH: 6370						
; TYPE: DNA						
; ORGANISM: Homo sapiens						
US-09-893-238-12						
Query Match 35.2%; Score 1510; DB 9; Length 6370;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 1510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	2781	GCAGTCCCGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCAACGCGCAGCTC	2840			
Db	548	GCAGTCCCGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCAACGCGCAGCTC	607			
QY	2841	TGAGTGCATGTGGTGCACACATGAACAGTGTGGACTCCCAATGCTATGTGGCTC	2900			
Db	608	TGAGTGCATGTGGTGCACACATGAACAGTGTGGACTCCCAATGCTATGTGGCTC	667			
QY	2901	CTTCCCTTTGGCCAGTGTATGGAAATGATACAGTACAGACCTGCCCTCCCTGAAATTTG	2960			
Db	668	CTTCCCTTTGGCCAGTGTATGGAAATGATACAGTACAGACCTGCCCTCCCTGAAATTTG	727			
QY	2961	TTGAGGTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTACTGA	3020			
Db	728	TTGAGGTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTACTGA	787			
QY	3021	TCCGAGCAATACTGCGAAGGGAATGCATAGAGGTTTCTTAAAGACCAAGTGAAGAT	3080			
Db	788	TCCGAGCAATACTGCGAAGGGAATGCATAGAGGTTTCTTAAAGACCAAGTGAAGAT	847			

Qy	4161	GTGAGTCTTTGCTGTGGCAGCGCCCTGGTGGACATTCTTCAGCAGATGCCGATAGTGTA	4220
Db	1928	GTGAGTCTTTGCTGTGGCAGCGCCCTGGTGGACATTCTTCAGCAGATGCCGATAGTGTA	1987
Qy	4221	CAAGGAGATCTAGGAGCCCTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGAC	4280
Db	1988	CAAGGAGATCTAGGAGCCCTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGAC	2047
Qy	4281	CTGCATCTGA	4290
Db	2048	CTGCATCTGA	2057

```

RESULT 5
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 429
; SOFTWARE: fastseq for windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-1

```

RESULT 6
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary

```

; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liguu
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match          5.7%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred No. 2.1e-118;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGCG 2574
Db 246 CTGCGAATATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGCG 187

QY 2575 CTTCCGAACATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCGCCATTTACAAATAGT 2634
Db 186 CTTCCGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCGCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCCATGTGTCGGATTCTGTGGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCCATGTGTCGGATTCTGTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

```

```

RESULT 7
US-09-902-941-966/c
/ Sequence 966, Application US/09902941
/ Patent NO. US20020172952A1
/ GENERAL INFORMATION:
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Watanabe, Yoshihiro
/ APPLICANT: Johnson, Jeffrey C.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Marnerakis, Margarita
/ APPLICANT: Carter, Darrick
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: McNabb, Andria
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C17
/ CURRENT APPLICATION NUMBER: US/09/902.941
/ CURRENT FILING DATE: 2001-07-10
/ NUMBER OF SEQ ID NOS: 2002
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 966
/ LENGTH: 246
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-902-941-966

```


Query Match 5.7%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.1e-118;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTCTCTACTGCTGGGAGATATGCCCCATTTACAAATAGT 2634
Db 186 CTTGCGAAGATCAATGTCTCTACTGCTGGGAGATATGCCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGAAATCTGTTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 8

US-09-849-626-966/c

; Sequence 966, Application US/09849626

; Publication No. US20020197669A1

; GENERAL INFORMATION:

; APPLICANT: Bangur, Chaitanya

; APPLICANT: Fanger, Gary

; APPLICANT: Wang, Aijun

; APPLICANT: Wang, Tongtong

; APPLICANT: Switzer, Anne

; APPLICANT: McNeill, Patricia

; APPLICANT: Clapper, Jonathan

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C16

; CURRENT APPLICATION NUMBER: US/09/849,626

; CURRENT FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 1926

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-849-626-966

Query Match 5.7%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.1e-118;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTCTCTACTGCTGGGAGATATGCCCCATTTACAAATAGT 2634
Db 186 CTTGCGAAGATCAATGTCTCTACTGCTGGGAGATATGCCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGAAATCTGTTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 9

US-10-017-754-966/c

; Sequence 966, Application US/10017754

; Publication No. US20030054363A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C18

; CURRENT APPLICATION NUMBER: US/10/017,754

; CURRENT FILING DATE: 2001-10-29

; NUMBER OF SEQ ID NOS: 2004

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-017-754-966

Query Match 5.7%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.1e-118;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTCTCTACTGCTGGGAGATATGCCCCATTTACAAATAGT 2634
Db 186 CTTGCGAAGATCAATGTCTCTACTGCTGGGAGATATGCCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGAAATCTGTTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGATGCTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 10

US-10-113-872-966/c

; Sequence 966, Application US/10113872

; Publication No. US20030170255A1

; GENERAL INFORMATION:

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C19

; CURRENT APPLICATION NUMBER: US/10/113,872

; CURRENT FILING DATE: 2002-03-28


```

;
; NAME/KEY: misc_feature
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

Query Match
Best Local Similarity 5.3%; Score 228; DB 14; Length 625;
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2699 GTACTCGGAGCTGAAGGCTGCAACTGTCATCAACCCCACTCAATGGTAGTGTCTGAAA 2758
DB 340 GTACTCGGAGCTGAAGGCTGCAACTGTCATCAACCCCACTCAATGGTAGTGTCTGAAA 281
QY 2759 GGCCTGCAAAACACAGTGTAAAGCAGTGCAGACACCATGTCCTTGAGGACAGCATGTG 2818
DB 280 GGCCTGCAAAACACAGTGTAAAGCAGTGCAGACACCATGTCCTTGAGGACAGCATGTG 221
QY 2819 GAGATTGCACAGCGCAGCTCTGAGTGCATGTTGGTGACGAAACATGAAGCAGTGTGTGG 2878
DB 220 GAGATTGCACAGCGCAGCTCTGAGTGCATGTTGGTGACGAAACATGAAGCAGTGTGTGG 161
QY 2879 ACTCCAATCCCTATGTGGCTCTCTCCCTTTGGCCAGTGTAAAGTGGTATGATGA 2938
DB 160 ACTCCAATCCCTATGTGGCTCTCTCCCTTTGGCCAGTGTAAAGTGGTATGATGA 101
QY 2939 GCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCT 2977
DB 100 GCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCT 62

RESULT 14
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

;
; NAME/KEY: misc_feature
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
; OTHER INFORMATION: n = A,T,C or G
US-09-864-761-15900

Query Match
Best Local Similarity 4.9%; Score 210; DB 9; Length 531;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1238 CAACTGGGAATGTGACCAATGAGTTCAGAGTTCATATTAATGATGATGGTGT 1297
DB 184 CAACTGGGAATGTGACCAATGAGTTCAGAGTTCATATTAATGATGATGGTGT 243
QY 1298 TGTGACCCCTAAGGCAAGGACCATATCCAGTGTGGGCACTCTGCACACATTGTTA 1357
DB 244 TGTGACCCCTAAGGCAAGGACCATATCCAGTGTGGGCACTCTGCACACATTGTTA 303
QY 1358 CACTGAAGAAATGCCGAGTGTGATGTCATCTTGTGTCATCTTGTGTCACCTCTCTATGGAT 1417
DB 304 CACTGAAGAAATGCCGAGTGTGATGTCATCTTGTGTCATCTTGTGTCACCTCTCTATGGAT 363
QY 1418 ATATAGCAATGTGCAGGAATATGATTTGG 1447
DB 364 ATATAGCAATGTGCAGGAATATGATTTGG 393

RESULT 15
US-09-864-761-32404
; Sequence 32404, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```


RESULT 17
US-09-864-761-4720/c
; Sequence 4720, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4720
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL13273.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
US-09-864-761-4720

Query Match

4.0%; Score 173; DB 9; Length 375;

Best Local Similarity 100.0%; Pred. No. 6.7e-80;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCTACTGCTGCTGGAGATATCTCCCAATTTACAAATAGTTTACTACAGTGGATGCC 2651
DB 375 GTCTACTGCTGCTGGAGATATCTCCCAATTTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTAGCCCCAGTGTCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTTCGGGACT 2711
DB 315 GTCTAGCCCCAGTGTCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTTCGGGACT 256
QY 2712 GAAGGTGCAACTGCATCAACCCACTCAATGCTAGTCTCTGGAAGCCCTG 2764
DB 255 GAAGGTGCAACTGCATCAACCCACTCAATGCTAGTCTCTGGAAGCCCTG 203

RESULT 18

US-09-864-761-3681/c
; Sequence 3681, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3681
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

OTHER INFORMATION: MAP TO AL132773.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
US-09-864-761-3681

Query Match 3.3%; Score 141; DB 9; Length 402;
Best Local Similarity 100.0%; Pred. No. 5e-63;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241
Db |||||
QY 391 CAGCTTGGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332
Db |||||
QY 3242 AGAAGCTGACACAGCAAGCACTCGAGAGCTGCATATCTGGTCTTACGGTGATCCCA 3301
Db |||||
QY 331 AGAAGCTGACACAGCAAGCACTCGAGAGCTGCATATCTGGTCTTACGGTGATCCCA 272
Db |||||
QY 3302 CCAATGGAGGGAATGTGAGC 3322
Db |||||
QY 271 CCAATGGAGGGAATGTGAGC 251
Db |||||

RESULT 19
US-09-796-692-7610
; Sequence 7610, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: unsure
LOCATION: (140)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (142)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (162)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (196)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (204)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (205)
OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7610

Query Match 3.2%; Score 139; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2913 CCAGTGTATGGAATGGTATACGATGAGCACCTGCCCTGAAATTTGTTACGGTACTG 2972
Db |||||
QY 2973 TACCTGTAGTCATCTGCTGGAGCAACAGGCTGTGGCTGTACTGATCCAGCAATAC 3032
Db |||||
QY 61 TACCTGTAGTCATCTGCTGGAGCAACAGGCTGTGGCTGTACTGATCCAGCAATAC 120
Db |||||
QY 3033 TGGCAAGGGAATGCATA 3051
Db |||||
QY 121 TGGCAAGGGAATGCATA 139
Db |||||

RESULT 20
US-10-040-862-7610
; Sequence 7610, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT FILING DATE: 2001-11-06
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03

	PRIOR APPLICATION NUMBER: US 60/223,416	
	PRIOR FILING DATE: 2000-08-04	
	PRIOR APPLICATION NUMBER: US 60/223,378	
	PRIOR FILING DATE: 2000-08-07	
	PRIOR APPLICATION NUMBER: US 09/796,692	
	PRIOR FILING DATE: 2001-03-01	
	NUMBER OF SEQ ID NOS: 10467	
	SOFTWARE: FastSeq for Windows Version 3.0	
	SEQ ID NO 7610	
	LENGTH: 212	
	TYPE: DNA	
	ORGANISM: Homo sapiens	
	FEATURE:	
	NAME/KEY: unsure	
	LOCATION: (140)	
	OTHER INFORMATION: n=A,T,C or G	
	FEATURE:	
	NAME/KEY: unsure	
	LOCATION: (142)	
	OTHER INFORMATION: n=A,T,C or G	
	FEATURE:	
	NAME/KEY: unsure	
	LOCATION: (162)	
	OTHER INFORMATION: n=A,T,C or G	
	FEATURE:	
	NAME/KEY: unsure	
	LOCATION: (196)	
	OTHER INFORMATION: n=A,T,C or G	
	FEATURE:	
	NAME/KEY: unsure	
	LOCATION: (204)	
	OTHER INFORMATION: n=A,T,C or G	
	FEATURE:	
	NAME/KEY: unsure	
	LOCATION: (205)	
	OTHER INFORMATION: n=A,T,C or G	
	US-10-040-862-7610	
	Query Match 3.2%; Score 139; DB 14; Length 212;	
	Best Local Similarity 100.0%; Pred. No. 5.5e-62;	
	Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	2913 CCAGTGTATGGAAATGTTACGATGAGCACCTGCCCCCCCTGAAAAATTGTTCAGGCTACTG	2972
Db	1 CCAGTGTATGGAAATGTTACGATGAGCACCTGCCCCCCCTGAAAAATTGTTCAGGCTACTG	60
QY	2973 TACTGTAGTCATTGCTTGAGGCAACCAGAGCTGTGGCTGGTGTACTGATCCCGACCAATAC	3032
Db	61 TACTGTAGTCATTGCTTGAGGCAACCAGAGCTGTGGCTGGTGTACTGATCCCGACCAATAC	120
QY	3033 TGGCAAAGGGAATGCATA 3051	
Db	121 TGGCAAAGGGAATGCATA 139	
	RESULT 21	
	US-10-057-475B-7610	
	; Sequence 7610, Application US/10057475B	
	; Publication No. US20040002068A1	
	; GENERAL INFORMATION:	
	; APPLICANT: Gaiger, Alexander	
	; APPLICANT: Algate, Paul A.	
	; APPLICANT: Mannion, Jane	
	; APPLICANT: Clapper, Jonathan David	
	; APPLICANT: Wang, Aijun	
	; APPLICANT: Ordonez, Nadia	
	; APPLICANT: Carter, Lauren	
	; APPLICANT: McNeill, Patricia Dianne	
	; APPLICANT: Corixa Corporation	
	; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy	
	; FILE REFERENCE: 014058-014402US	
	; CURRENT APPLICATION NUMBER: US/10/057,475B	

```

; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(212)
; OTHER INFORMATION: n = g, a, c o r t
US-10-057-475B-7610

Query Match 3.2%; Score 139; DB 15; Length 212;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2913 CCAGTGATATGGAATGTTATACGATGAGCACCTGCCCCCTGAAAAATGTTTCAGGCTACTG 2972
Db 1 CCAGTGATATGGAATGTTATACGATGAGCACCTGCCCCCTGAAAAATGTTTCAGGCTACTG 60

QY 2973 TACCTGTAGTCATTGCTTGGAGCAACGAGGCTGTGGTGTACTGTATCCAGCAATAC 3032
Db 61 TACCTGTAGTCATTGCTTGGAGCAACGAGGCTGTGGTGTACTGTATCCAGCAATAC 120

QY 3033 TGGCAAGGGAAATGCATA 3051
Db 121 TGGCAAGGGAAATGCATA 139

RESULT 22
US-10-154-884B-7610
; Sequence 7610, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303

```



```
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(212)
; OTHER INFORMATION: n = g, a, c or t
; US-10-154-884B-7610

Query Match          3.2%; Score 139; DB 15; Length 212;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2913 CCAGTGTATGAATGGTATACATGAGCACCTGCCCTGAAATTTCTCAGGCTACTG 2972
Db 1 CCAAGTGTATGAATGGTATACATGAGCACCTGCCCTGAAATTTCTCAGGCTACTG 60

QY 2973 TACCTGTAGTCAATGCTGGACCAACCAAGCTGTGGCTGGTGTACTGTATCCAGCAATAC 3032
Db 61 TACCTGTAGTCAATGCTGGACCAACCAAGCTGTGGCTGGTGTACTGTATCCAGCAATAC 120

QY 3033 TGGCAAGGAAATGCATA 3051
Db 121 TGGCAAGGAAATGCATA 139

RESULT 23
US-09-864-761-20449/c
; Sequence 20449, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20449
; LENGTH: 137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EST HUMAN HIT: BF672370.1, EVALUATE 2.00e-71
; OTHER INFORMATION: NT HIT: AB011120.1, EVALUATE 1.00e-71
; OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUATE 3.00e-11
; US-09-864-761-20449

Query Match          3.2%; Score 137; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 6.1e-61;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3185 CTTGCCAATGCAACGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAAGTGTGAGA 3244
Db 137 CTTGCCAATGCAACGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAAGTGTGAGA 78

QY 3245 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3304
Db 77 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 18

QY 3305 ATGCGAGGAAATGTCTAG 3321
Db 17 ATGCGAGGAAATGTCTAG 1

RESULT 24
US-10-027-632-264639
; Sequence 264639, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```

;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 264639
;; LENGTH: 594
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-264639

Query Match 2.0%; Score 86; DB 15; Length 594;
Best Local Similarity 100.0%; Pred. No. 5.1e-34;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1786 GCGTGTGACCGTGTGTCAGTGTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1845
DB 144 GCGTGTGACCGTGTGTCAGTGTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 203

QY 1846 GCGCATTTCAGCAGTCTTACACACAG 1871
DB 204 GCGCATTTCAGCAGTCTTACACACAG 229

RESULT 25
US-10-027-632-264640
;; Sequence 264640, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 264640
;; LENGTH: 594
;; TYPE: DNA
;; ORGANISM: Human

US-10-027-632-264640

Query Match 2.0%; Score 86; DB 15; Length 594;
Best Local Similarity 100.0%; Pred. No. 5.1e-34;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1786 GCGTGTGACCGTGTGTCAGTGTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1845
DB 144 GCGTGTGACCGTGTGTCAGTGTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 203

QY 1846 GCGCATTTCAGCAGTCTTACACACAG 1871
DB 204 GCGCATTTCAGCAGTCTTACACACAG 229

RESULT 26

US-10-027-632-270240
;; Sequence 270240, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 270240
;; LENGTH: 629
;; TYPE: DNA
;; ORGANISM: Human

US-10-027-632-270240

Query Match 1.8%; Score 79; DB 15; Length 629;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3565 CAAAACAGGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACC 3624
DB 347 CAAAACAGGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACC 406

QY 3625 TGGGCTGCCAGTTTCTCAG 3643
DB 407 TGGGCTGCCAGTTTCTCAG 425

RESULT 27

US-09-908-975-15421
;; Sequence 15421, Application US/09908975
;; Publication No. US20030165843A1
;; GENERAL INFORMATION:
;; APPLICANT: SHOSHAN, Avi
;; APPLICANT: WASSERMAN, Alon
;; APPLICANT: MINTZ, Eli
;; APPLICANT: FAIGLER, Simchon
;; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;; FILE REFERENCE: 36688-0005
;; CURRENT APPLICATION NUMBER: US/09/908,975
;; PRIOR FILING DATE: 2001-07-20
;; PRIOR APPLICATION NUMBER: US 60/287,724
;; PRIOR FILING DATE: 2001-05-02
;; PRIOR APPLICATION NUMBER: US 60/221,607
;; PRIOR FILING DATE: 2000-07-28
;; NUMBER OF SEQ ID NOS: 32337
;; SOFTWARE: PatentIn version 3.0

US-09-908-975-15421

```
; SEQ ID NO 15421
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-15421

Query Match
1.4%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3e-20; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 0;

QY 3547 GTGGTACTCTCGAGCAAAACAGGGATTGGACATGTTTCATCAATGCCCTCCAAGAAAT 3606
Db 1 GTGGTACTCTCTCGAGCAAAACAGGGATTGGACATGTTTCATCAATGCCCTCCAAGAAAT 60

RESULT 28
US-09-893-238-5
; Sequence 5, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-5

Query Match
1.3%; Score 55; DB 9; Length 90050;
Best Local Similarity 100.0%; Pred. No. 1.4e-17; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 0;

QY 3130 TCCAGCATGTGCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCACATGTCACG 3184
Db 3184 TCCAGCATGTGCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCACATGTCACG 31894

RESULT 29
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
```

```
; ORGANISM: Mus musculus
US-09-893-238-10

Query Match
1.2%; Score 53; DB 9; Length 1051;
Best Local Similarity 100.0%; Pred. No. 1.3e-16; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 0;

QY 367 GGCTGGTGGCGGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGG 419
Db 571 GGCTGGTGGCGGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGG 623

RESULT 30
US-09-893-238-8
; Sequence 8, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-8

Query Match
1.1%; Score 47; DB 9; Length 2419;
Best Local Similarity 100.0%; Pred. No. 2e-13; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;

QY 1624 CAGATCTGACCATCTTTAAGGACAGCCGATTTTCGGTACTTGA 1670
Db 1890 CAGATCTGACCATCTTTAAGGACAGCCGATTTTCGGTACTTGA 1936

RESULT 31
US-10-027-632-108988/c
; Sequence 108988, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108988
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(468)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-108988

Query Match
Best Local Similarity 0.9%; Score 38; DB 15; Length 468;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3417 ATGTGAGTACGAGTATCGATACCAAGGAAACCCCTCTCA 3454
DB 414 ATGTGAGTACGAGTATCGATACCAAGGAAACCCCTCTCA 377

RESULT 32
US-09-893-238-4
; Sequence 4, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 5973
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-4

Query Match
Best Local Similarity 0.8%; Score 35; DB 9; Length 5973;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 TGGGACCAATTATATGTTTATGATGGGACTCAAT 581
DB 4379 TGGGACCAATTATATGTTTATGATGGGACTCAAT 4413

RESULT 33
US-09-908-975-29990
; Sequence 29990, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchaon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607

```

```

; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29990
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-29990

Query Match
Best Local Similarity 0.7%; Score 32; DB 10; Length 65;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2860 AACATGAAGCAGTGTGTGGACTCCAATGCTA 2891
DB 34 AACATGAAGCAGTGTGTGGACTCCAATGCTA 65

RESULT 34
US-09-893-238-3/c
; Sequence 3, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-3

Query Match
Best Local Similarity 0.7%; Score 32; DB 9; Length 17056;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 TGTGACATTCTCCTCAGTACAGACAAGTGTGG 875
DB 3978 TGTGACATTCTCCTCAGTACAGACAAGTGTGG 3947

RESULT 35
US-09-893-238-23/c
; Sequence 23, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 23
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-893-238-23

Query Match 0.7%; Score 29; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 552 CCATTATATATGTTATGATGGGACTCAA 590
Db 29 CCATTATATGTTATGATGGGACTCAA 1

RESULT 36

US-10-419-296-1
; Sequence 1, Application US/10419296
; Publication No. US20040014141A1
; GENERAL INFORMATION:
; APPLICANT: Woolf, Clifford J.
; APPLICANT: Samad, Tarek A.
; TITLE OF INVENTION: DRG11-RESPONSIVE (DRAGON) GENE FAMILY
; FILE REFERENCE: 00786/419002
; CURRENT APPLICATION NUMBER: US/10/419,296
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: 60/373,519
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-419-296-1

Query Match 0.7%; Score 29; DB 15; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 CCCGCCGCGCTGTGCTGCTGCTGCTGC 232
Db 114 CCCGCCGCGCTGTGCTGCTGCTGCTGC 142

RESULT 37

US-10-029-386-22666
; Sequence 22666, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22666
; LENGTH: 1622
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133445.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q05175, EVALUATE 4.90e+00

; OTHER INFORMATION: NT HIT: g115595198, EVALUATE 1.60e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA700990.1, EVALUATE 0.00e+00
US-10-029-386-22666

Query Match 0.7%; Score 29; DB 14; Length 1622;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 GGCCGAGCGCGCGCGCGCGCGCGCGG 268
Db 967 GGCCGAGCGCGCGCGCGCGCGCGCGG 995

RESULT 38

US-10-311-455-169
; Sequence 169, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: Olek, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determin
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 169
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-169

Query Match 0.6%; Score 27; DB 14; Length 5312;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 CGCGCGCGCGCGCGCGCGCGGTGCGGG 275
Db 5249 CGCGCGCGCGCGCGCGCGCGGTGCGGG 5275

RESULT 39

US-09-893-238-48
; Sequence 48, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 24
; TYPE: DNA

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-893-238-48

Query Match          0.6%; Score 24; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1631 GGACCATTCCTTAAGCAGCGCGAT 1654
Db 1 GGACCATTCCTTAAGCAGCGCGAT 24

RESULT 40
US-09-783-590-7412
; Sequence 7412, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7412
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (141)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (152)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (271)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (319)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (325)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (332)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (380)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
```

```
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (388)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (390)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (393)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (395)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (397)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (420)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (444)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (448)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (452)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (463)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (471)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (483)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (491)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (493)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (497)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-7412
```

```
Query Match          0.6%; Score 24; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 205 CCGCCGCCGCTGTTGCTGCTG 228
|||
Db 164 CCGCCGCCGCTGTTGCTGCTG 187

RESULT 41

```

US-10-424-599-106383/C
; Sequence 106383, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 106383
; LENGTH: 784
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67080C.1
US-10-424-599-106383

```

Query Match 0.6%; Score 24; DB 12; Length 784;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels

Qy 180 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 203
|||
db 588 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565

RESULT 42

```

US-10-425-114-13318/C
; Sequence 13318, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Tongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13318
; LENGTH: 899
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMLEC1810087E09_FLI
US-10-425-114-13318

```

Query Match 0.6%; Score 24; DB 12; Length 899;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels

Qy 180 GCTGCTGCTGCTGCTGTTGCT 203
|||||
Ddb 341 GCTGCTGCTGCTGCTGTTGCT 318

RESULT 43

```

US-10-425-114-29183/c
; Sequence 29183, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29183
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB83-013-H1_FLI
US-10-425-114-29183

```

Query Match 0.6%; Score 24; DB 12; Length 1003;
Best Local Similarity 100.0%; Pred.No. 0.26;
Matches 24; Conservative 0; Mismatches 0; Indels

Qy 247 GCCGCGCGCGCGCGCGCGTG 270
Db 595 GCCGCGCGCGCGCGCGCGTG 572

RESULT 44

```

US-10-424-599-10899/c
/ Sequence 10899, Application US/10424599
/ Publication NO. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223) B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 10899
/ LENGTH: 1097
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_109851C.1
US-10-424-599-10899

```

```
Query Match      0.6%; Score 24; DB 12; Length 1097;
Best Local Similarity 100.0%; Pred.No. 0.26;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 180 GCTGCTGCTGCTGCTGCTGCT 203
|||
db 512 GCTGCTGCTGCTGCTGCTGCT 489

RESULT 45

US-10-424-599-547/c
; Sequence 547, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 547
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1372)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100494C.1
US-10-424-599-547

Query Match 0.6%; Score 24; DB 12; Length 1372;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 529 GCTGCTGCTGCTGCTGCTGCTGCT 506

RESULT 46
US-10-424-599-21459/c
; Sequence 21459, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 21459
; LENGTH: 1756
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_119382C.1
US-10-424-599-21459

Query Match 0.6%; Score 24; DB 12; Length 1756;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 1198 GCTGCTGCTGCTGCTGCTGCTGCT 1175

RESULT 47
US-09-814-353-21391/c
; Sequence 21391, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353

CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21391
; LENGTH: 2180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 45, 2179, 2180
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21391

Query Match 0.6%; Score 24; DB 10; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 1050 GCTGCTGCTGCTGCTGCTGCTGCT 1027

RESULT 48
US-10-437-171-3/c
; Sequence 3, Application US/10437171
; Publication No. US20030235564A1
; GENERAL INFORMATION:
; APPLICANT: Doll, Bruce
; APPLICANT: Fu, Huihua
; APPLICANT: Hollinger, Jeffrey O.
; APPLICANT: Slier, Charles
; TITLE OF INVENTION: Compositions and Devices Comprising or Encoding the Run X2
; FILE REFERENCE: 1915/14014US02
; CURRENT APPLICATION NUMBER: US/10/437,171
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/380,554
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; NAME/KEY: CDS
; LOCATION: (1)..(1644)
US-10-437-171-3

Query Match 0.6%; Score 24; DB 15; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 223 GCTGCTGCTGCTGCTGCTGCTGCT 200

RESULT 49

US-10-437-171-1/c
; Sequence 1, Application US/10437171
; Publication No. US20030235564A1
; GENERAL INFORMATION:
; APPLICANT: Doll, Bruce
; APPLICANT: Fu, Huibua
; APPLICANT: Hollinger, Jeffrey O.
; APPLICANT: Siet, Charles
; TITLE OF INVENTION: Compositions and Devices Comprising or Encoding the Run X2
; TITLE OF INVENTION: Protein and Method of Use
; FILE REFERENCE: 1915/14014US02
; CURRENT APPLICATION NUMBER: US/10/437,171
; PRIOR FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/380,554
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Primer
US-10-437-171-1
Query Match 0.6%; Score 24; DB 15; Length 3334;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 531 GCTGCTGCTGCTGCTGCTGCTGCT 508
RESULT 50
US-10-424-599-56004/c
; Sequence 56004, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285664
; SEQ ID NO 56004
; LENGTH: 4964
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_21582C.1
US-10-424-599-56004
Query Match 0.6%; Score 24; DB 12; Length 4964;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 405 GCTGCTGCTGCTGCTGCTGCTGCT 382
RESULT 51
US-10-198-846-9854/c
; Sequence 9854, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: PasSeq for Windows Version 4.0
; SEQ ID NO 9854
; LENGTH: 5085
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 5070, 5071, 5072, 5073, 5074, 5075, 5076, 5077, 5078,
; LOCATION: 5079, 5080, 5081, 5082, 5083, 5084, 5085
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9854

Query Match 0.6%; Score 24; DB 14; Length 5085;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 1432 GCTGCTGCTGCTGCTGCTGCTGCT 1409

RESULT 52
US-09-819-104A-6/c
; Sequence 6, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 7386
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(7386)
US-09-819-104A-6
Query Match 0.6%; Score 24; DB 10; Length 7386;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
DB 1519 GCTGCTGCTGCTGCTGCTGCTGCT 1496

RESULT 53
US-09-819-104A-3/c
; Sequence 3, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR

```
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(7521)
US-09-819-104A-3

Query Match      0.6%; Score 24; DB 10; Length 7521;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1504 GCTGCTGCTGCTGCTGCTGCTGCT 1481

RESULT 54
US-10-414-692-14/c
; Sequence 14, Application US/10414692
; Publication No. US20030228607A1
; GENERAL INFORMATION:
; APPLICANT: X-Ceptor Therapeutics, Inc.
; TITLE OF INVENTION: Screening method and modulators having an improved therapeutic
; FILE REFERENCE: 8012-002-US
; CURRENT APPLICATION NUMBER: US/10/414,692
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/372,650
; PRIOR FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-414-692-14

Query Match      0.6%; Score 24; DB 15; Length 7554;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1507 GCTGCTGCTGCTGCTGCTGCTGCT 1484

RESULT 55
US-09-819-104A-4/c
; Sequence 4, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 8544
; TYPE: DNA
```

```
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(7545)
US-09-819-104A-4

Query Match      0.6%; Score 24; DB 10; Length 8544;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1678 GCTGCTGCTGCTGCTGCTGCTGCT 1655

RESULT 56
US-10-174-014-11/c
; Sequence 11, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
; FILE REFERENCE: PTS-0012
; CURRENT APPLICATION NUMBER: US/10/174,014
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 11
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(7555)
US-10-174-014-11

Query Match      0.6%; Score 24; DB 15; Length 8561;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1508 GCTGCTGCTGCTGCTGCTGCTGCT 1485

RESULT 57
US-09-819-104A-1/c
; Sequence 1, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)..(7677)
US-09-819-104A-1

Query Match      0.6%; Score 24; DB 10; Length 8686;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1660 GCTGCTGCTGCTGCTGCTGCTGCTGCT 1637

RESULT 58
US-10-174-014-4/c
; Sequence 4, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
; FILE REFERENCE: PIS-0012
; CURRENT APPLICATION NUMBER: US/10/174,014
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 4
; LENGTH: 8686
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)...(7680)
US-10-174-014-4

Query Match 0.6%; Score 24; DB 15; Length 8686;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1660 GCTGCTGCTGCTGCTGCTGCTGCTGCT 1637

RESULT 59
US-10-051-874-55/c
; Sequence 55, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zertusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie

; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595.
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 14859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-051-874-55

Query Match 0.6%; Score 24; DB 15; Length 14859;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203
Db 9082 GCTGCTGCTGCTGCTGCTGCTGCTGCT 9059

RESULT 60
US-10-085-198-111/c
; Sequence 111, Application US/10085198
; Publication No. US20040009907A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695

LOCATION: 77993
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 77994
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 77995
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 77996
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 77997
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 77998
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 77999
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78000
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78001
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78002
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78003
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78004
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78005
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78006
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78007
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78008
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78009
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78010
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78011

OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 78012
OTHER INFORMATION: unknown
FEATURE:
Query Match 0.6%; Score 24; DB 15; Length 221000;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 137744 GCTGCTGCTGCTGCTGCTGCTGCT 137721
RESULT 62
US-09-893-238-24
Sequence 24, Application US/09893238
Patent No. US20020150973A1
GENERAL INFORMATION:
APPLICANT: Moore, K.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
FILE REFERENCE: 7853-237
CURRENT APPLICATION NUMBER: US/09/893,238
CURRENT FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: 09/245,041
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/093,630
PRIOR FILING DATE: 1998-07-21
PRIOR APPLICATION NUMBER: 60/104,978
PRIOR FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 129
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 29
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Primer
US-09-893-238-24
Query Match 0.5%; Score 23; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 495 GCCAATAGATAATGAGACTTC 517
Db 7 GCCAATAGATAATGAGACTTC 29
RESULT 63
US-10-260-238-1970
Sequence 1970, Application US/10260238
Publication No. US20040016025A1
GENERAL INFORMATION:
APPLICANT: Budworth, Paul R.
APPLICANT: Moughamer, Todd G.
APPLICANT: Briggs, Steven P.
APPLICANT: Cooper, Bret
APPLICANT: Glazebrook, Jane
APPLICANT: Goff, Stephen A.
APPLICANT: Katagiri, Fumiyaki
APPLICANT: Kreps, Joel
APPLICANT: Provart, Nicholas
APPLICANT: Ricke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
FILE REFERENCE: 60111-NP
CURRENT APPLICATION NUMBER: US/10/260,238
CURRENT FILING DATE: 2002-09-26

; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1970
; LENGTH: 104
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (11)..(11)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (79)..(79)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-1970

Query Match 0.5%; Score 23; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 GCCGCGCGCGCGCGCGCGGT 269
|||
DB 80 GCCGCGCGCGCGCGCGCGGT 102
|||

RESULT 64

US-09-893-238-7/c
; Sequence 7, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238

; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-7

Query Match 0.5%; Score 23; DB 9; Length 221;
Best Local Similarity 100.0%; Pred. No. 0.8; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2713 AAGGCTGCAACCTGCATCAACCC 2735
|||
DB 118 AAGGCTGCAACCTGCATCAACCC 96
|||

RESULT 65

US-09-893-238-6/c
; Sequence 6, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-6

Query Match 0.5%; Score 23; DB 9; Length 226;
Best Local Similarity 100.0%; Pred. No. 0.8; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2713 AAGGCTGCAACCTGCATCAACCC 2735
|||
DB 123 AAGGCTGCAACCTGCATCAACCC 101
|||

RESULT 66

US-10-029-386-11303
; Sequence 11303, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.

; APPLICANT: Hanzel, David K.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11303
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: MAP TO AL022395.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
; OTHER INFORMATION: NT HIT: g113643518, EVALUE 1.00e-118
; OTHER INFORMATION: SWISSPROT HIT: P20285, EVALUE 1.00e-37
; OTHER INFORMATION: EST_HUMAN HIT: BF515537.1, EVALUE 1.00e-115
US-10-029-386-11303

Query Match 0.5%; Score 23; DB 14; Length 500;
Best Local Similarity 100.0%; Pred. No. 0.83; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GCCGCTGTTGCTGCTGCTGCTGC 232
|||
DB 342 GCCGCTGTTGCTGCTGCTGCTGC 364
|||

RESULT 67

US-10-260-238-5694/c
; Sequence 5694, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.

APPLICANT: Cooper, Bret
APPLICANT: Glazebrook, Jane
APPLICANT: Goff, Stephen A.
APPLICANT: Katagiri, Fumiyaki
APPLICANT: Kreps, Joel
APPLICANT: Provart, Nicholas
APPLICANT: Ricke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
FILE REFERENCE: 60111-NP
CURRENT APPLICATION NUMBER: US/10/260,238
CURRENT FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/325,448
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/325,277
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/370,620
PRIOR FILING DATE: 2002-04-04
NUMBER OF SEQ ID NOS: 6077
SEQ ID NO 5694
LENGTH: 510
TYPE: DNA
ORGANISM: Zea mays
US-10-260-238-5694

Query Match 0.5%; Score 23; DB 15; Length 510;
Best Local Similarity 100.0%; Pred.No. 0.63;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGTG 270
DB 49 CCGCGCGCGCGCGCGCGCGTG 27

RESULT 68
US-09-864-761-19977/c
Sequence 19977, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 19977
LENGTH: 625
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004827.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.7
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
OTHER INFORMATION: EST HUMAN HIT: AW881551.1, EVALUE 6.00e-31
OTHER INFORMATION: SWISSPROT HIT: QC4891, EVALUE 5.40e+00
OTHER INFORMATION: NT HIT: AF286598.1, EVALUE 0.00e+00
US-09-864-761-19977

Query Match 0.5%; Score 23; DB 9; Length 625;
Best Local Similarity 100.0%; Pred.No. 0.84;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GGCTGCTGCTGCTGCTGTTG 201
DB 429 GGCTGCTGCTGCTGTTG 407

RESULT 69
US-10-424-599-110798/c
Sequence 110798, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 110798
LENGTH: 728
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(728)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_71060C.1
US-10-424-599-110798

Query Match 0.5%; Score 23; DB 12; Length 728;
Best Local Similarity 100.0%; Pred.No. 0.85;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGGCGGCGGCGGCGGTGC 272
|||||
Db 207 GCGGCGGCGGCGGCGGCGGTGC 185

RESULT 70

US-10-425-114-20706
; Sequence 20706, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20706
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3245-068-H12_FLI
US-10-425-114-20706

Query Match 0.5%; Score 23; DB 12; Length 900;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 GCGCGCGCGCGCGCGCGCGCG 268
|||||
Db 540 GCGCGCGCGCGCGCGCGCGCG 562

RESULT 71

US-10-260-238-2125
; Sequence 2125, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; PRIORITY FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 2125
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N_region

; LOCATION: (4)..(4)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (6)..(7)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (769)..(769)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (857)..(858)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (930)..(930)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (1042)..(1042)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (1044)..(1044)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-2125

Query Match 0.5%; Score 23; DB 15; Length 1320;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGTC 202
|||||
Db 1140 GCTGCTGCTGCTGCTGCTGTC 1162

RESULT 72

US-09-864-761-3198/c
; Sequence 3198, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acemica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIORITY FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3198
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004827.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
US-09-864-761-3198

Query Match          0.5%; Score 23; DB 9; Length 1979;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GCGTCTGCTGCTGCTGCTGTTGTTG 201
    |||||
DB 733 GCGTCTGCTGCTGCTGTTGTTG 711

RESULT 73
US-10-260-238-1864/c
; Sequence 1864, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Rumiya
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1864
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (489)..(489)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (501)..(501)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (549)..(549)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (552)..(552)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (554)..(554)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region

Query Match          0.5%; Score 23; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGGTTG 270
    |||||
DB 1921 CCGCGCGCGCGCGCGGTTG 1899

RESULT 74
US-10-260-238-1920/c
; Sequence 1920, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Rumiya
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1920
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (489)..(489)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (501)..(501)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (549)..(549)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (552)..(552)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (554)..(554)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
```

```

; LOCATION: (559)...(559)
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1235)...(1238)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1281)...(1281)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1300)...(1300)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-1920

```

```

Query Match      0.5%; Score 23; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 250 GCGGCGGCGGCGGCGGCGTGC 272
Db 307 GCGGCGGCGGCGGCGGCGTGC 285

```

```

RESULT 75
US-10-108-260A-790/c
; Sequence 790, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 790
; LENGTH: 2264
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-790

```

```

Query Match      0.5%; Score 23; DB 15; Length 2264;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 210 GCGGCTGTGTGCTGCTGCTGC 232
Db 564 GCGGCTGTGTGCTGCTGCTGC 542

```

```

Search completed: March 1, 2004, 21:14:28
Job time : 1419 secs

```

This Page Blank (uspto)

ALIGNMENTS

RESULT 1
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14

Query Match 88.0%; Score 3774; DB 3; Length 8589;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3974; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	313	TGTGTCAACGGCGTCTGCAACCTGGCCAGTGGCTGCTGCGCCGCGCTGG	372
DB	310	TGTGTCAACGGCGTCTGCAACCTGGCCAGTGGCTGCTGCGCCGCGCTGG	369
QY	373	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAATGATCTTGGGTTT	432
DB	370	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAATGATCTTGGGTTT	429
QY	433	GTGACAGATGACCTGGAAATTAATAACAACGAGTGCAGTGGCTCATTAAGGA	492
DB	430	GTGACAGATGACCTGGAAATTAATAACAACGAGTGCAGTGGCTCATTAAGGA	489
QY	493	CAGCCAAATAGAAATATGAGACTTGGTTCAATCATTTTGTCTACAGAGTGTAGTGGAC	552
DB	490	CAGCCAAATAGAAATATGAGACTTGGTTCAATCATTTTGTCTACAGAGTGTAGTGGAC	549
QY	553	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTATTTAGTGGC	612
DB	550	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTATTTAGTGGC	609
QY	613	CTCATTTCTCGAGAGATGGCAATGAGACTGTCCCTGAGGTGTGTCACATCAGGT	672
DB	610	CTCATTTCTCGAGAGATGGCAATGAGACTGTCCCTGAGGTGTGTCACATCAGGT	669
QY	673	TATGCTTCTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAAT	732
DB	670	TATGCTTCTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAAT	729
QY	733	TACAGTTTGTATATGTGCCAAATTAACCTGTCAGGCCGAGGAGAGTGTAAAGATCAGTAAT	792
DB	730	TACAGTTTGTATATGTGCCAAATTAACCTGTCAGGCCGAGGAGAGTGTAAAGATCAGTAAT	789
QY	793	AGCAGCGAAATCTGTTGAATGCAATGTTCTGAAATCTGGAAGGTGAAGCATGTGACATT	852
DB	790	AGCAGCGAATCTGTTGAATGCAATGTTCTGAAATCTGGAAGGTGAAGCATGTGACATT	849
QY	853	CCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCATCTGCAATTTCAAGTGTATGTC	912
DB	850	CCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCATCTGCAATTTCAAGTGTATGTC	909

QY	913	AGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTTCTAGTTCTCTTACAGCT	972
DB	910	AGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTTCTAGTTCTCTTACAGCT	969
QY	973	AACCACTATTTTGGACTCGAGAGAAATTAATTAATTAAGCTCCCGAGGACATCTCAT	1032
DB	970	AACCACTATTTTGGACTCGAGAGAAATTAATTAATTAAGCTCCCGAGGACATCTCAT	1029
QY	1033	AAAGCTGTGTCATTAAGGAAACATTAATTAATTAAGCTCCCGAGGACATCTCAT	1092
DB	1030	AAAGCTGTGTCATTAAGGAAACATTAATTAATTAAGCTCCCGAGGACATCTCAT	1089
QY	1093	GATTATPAACATGTTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGT	1152
DB	1090	GATTATPAACATGTTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGT	1149
QY	1153	TCTGTGAAACAATGTGGTTGTTAGATATGTCATTTCTTTGGCATTATACAAGGATAAATTT	1212
DB	1150	TCTGTGAAACAATGTGGTTGTTAGATATGTCATTTCTTTGGCATTATACAAGGATAAATTT	1209
QY	1213	TACATGATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT	1272
DB	1210	TACATGATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT	1269
QY	1273	CACATTCATAATGAGTATGCTGGTGTGTTGACCCCTTAAGGCAAGGAGGACGATGTCAGTG	1332
DB	1270	CACATTCATAATGAGTATGCTGGTGTGTTGACCCCTTAAGGCAAGGAGGACGATGTCAGTG	1329
QY	1333	GTGGGCACTCTGCAACATTTGTACACTGAAGAATGGCCAGTGGTCTATGCTGGTCAATC	1392
DB	1330	GTGGGCACTCTGCAACATTTGTACACTGAAGAATGGCCAGTGGTCTATGCTGGTCAATC	1389
QY	1393	TTTGTGCTACCTGCTCTCTATGATATATAAGCAATGTGCAAGGATATGATTTGGATAG	1452
DB	1390	TTTGTGCTACCTGCTCTCTATGATATATAAGCAATGTGCAAGGATATGATTTGGATAG	1449
QY	1453	AAACATGAGATATATTAACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC	1512
DB	1450	AAACATGAGATATATTAACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC	1509
QY	1513	AGTGTTTACGACCATAGGACCCAGGCTCTATACGTTTCATGTGGTGTGCTACAAGGCTTTT	1572
DB	1510	AGTGTTTACGACCATAGGACCCAGGCTCTATACGTTTCATGTGGTGTGCTACAAGGCTTTT	1569
QY	1573	GCCAAATAGTACCGGCTTGCAGATGATCTCTACCATATGATGTCGATACCCAGATGTCG	1632
DB	1570	GCCAAATAGTACCGGCTTGCAGATGATCTCTACCATATGATGTCGATACCCAGATGTCG	1629
QY	1633	ACCATTCTTTAAGGACAGCCGATTTTTCGGTTTACTTTGCAACAGCTGTGATAGTGTGGA	1692
DB	1630	ACCATTCTTTAAGGACAGCCGATTTTTCGGTTTACTTTGCAACAGCTGTGATAGTGTGGA	1689
QY	1693	ACCATGCTGGTGTGGGGGAAACACACAAATGACATCTATGAGCCATGCGGCCAAA	1752
DB	1690	ACCATGCTGGTGTGGGGGAAACACACAAATGACATCTATGAGCCATGCGGCCAAA	1749
QY	1753	TGCTTCTCTTCAGATTTTCAATGGCTTATGACATTTGCTGTGACCGCTGTGCTGCTTCCC	1812
DB	1750	TGCTTCTCTTCAGATTTTCAATGGCTTATGACATTTGCTGTGACCGCTGTGCTGCTTCCC	1809
QY	1813	AGACCTGATCTCCACATGATGTCACAGATTTGGCCATTTAGCAGTCTTACACAAACAGC	1872
DB	1810	AGACCTGATCTCCACATGATGTCACAGATTTGGCCATTTAGCAGTCTTACACAAACAGC	1869
QY	1873	ACCATGATGCTGCTGGTGTGTTCAATAGTCTCTCTCAGCAGCATCTCTGTTATTCACC	1932
DB	1870	ACCATGATGCTGCTGGTGTGTTCAATAGTCTCTCTCAGCAGCATCTCTGTTATTCACC	1929
QY	1933	TCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTCTTTAGCAGCAGGACCTGGTAT	1992
DB	1930	TCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTCTTTAGCAGCAGGACCTGGTAT	1989
QY	1993	CGGTGTGTGTGAAACACAGGGGTCTCTCAGTGTATCTCTGCTGGGCTGCGCACTGATGAA	2052

Db 1990 CGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGTTGGGCGCTGGCAACTGATGAA 2049
Qy 2053 CAGAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGT 2112
Db 2050 CAAGNAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGT 2109
Qy 2113 GACGACACACAGATGTGTTACAGTGTATACAGCCAAACCAATGACTGCGCACTGTGCAAT 2172
Db 2110 GACCAGCACACAGATGTTTACAGTGTGACAGCCAAACCAATGACTGCGCACTGTGCAAT 2169
Qy 2173 GACCATTTGTCGCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCATTTTGTAGTAT 2232
Db 2170 GACCATTTGTCGCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCATTTTGTAGTAT 2229
Qy 2233 GAGAAATGCCCCAAAGGATAACCCATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2292
Db 2230 GAGAAATGCCCCAAAGGATAACCCATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2289
Qy 2293 TGTGCCCTGGACAGAACTGCGCAGTGGAGAGGCCCGGAAATCAGGAGTGCAATGCGCCTGCC 2352
Db 2290 TGTGCCCTGGACAGAACTGCGCAGTGGAGAGGCCCGGAAATCAGGAGTGCAATGCGCCTGCC 2349
Qy 2353 GAAATAATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCATGTTTGAAATTTACTACT 2412
Db 2350 GAAATAATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCATGTTTGAAATTTACTACT 2409
Qy 2413 GCCAAGAGAAATATGACAAATGCTTAAATGTTTGTAGGAAACCAATGCGCCTTTGGCT 2472
Db 2410 GCCAAGAGAAATATGACAAATGCTTAAATGTTTGTAGGAAACCAATGCGCCTTTGGCT 2469
Qy 2473 TCTCTTACACCCAGAGAGAGTGAATTTGCTTAAAGCAGCTGCGAATAATGCACTCA 2532
Db 2470 TCTCTTACACCCAGAGAGAGTGAATTTGCTTAAAGCAGCTGCGAATAATGCACTCA 2529
Qy 2533 TCTCAGACATGTCCAGCTCACTTAAACCCATGGGTGCGCCTTGGAGATCAATGTG 2592
Db 2530 TCTCAGACATGTCCAGCTCACTTAAACCCATGGGTGCGCCTTGGAGATCAATGTG 2589
Qy 2593 TCCTACTGGTCTGGAGATATGCCCCATTTACAAATGTTTACTACAGTGATGCGG 2652
Db 2590 TCCTACTGGTCTGGAGATATGCCCCATTTACAAATGTTTACTACAGTGATGCGG 2649
Qy 2653 TCTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 2712
Db 2650 TCTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 2709
Qy 2713 AAGGCTGCAACTGTGATCAACCCACTCAATGATGTCTGTGAAGCCCTGCAACAC 2772
Db 2710 AAGGCTGCAACTGTGATCAACCCACTCAATGATGTCTGTGAAGCCCTGCAACAC 2769
Qy 2773 AGTGCTAAGCAGTGCAGACACCATGTCCTTGAAGACAGATGTGAGATTTGACACAGC 2832
Db 2770 AGTGCTAAGCAGTGCAGACACCATGTCCTTGAAGACAGATGTGAGATTTGACACAGC 2829
Qy 2833 GGCAGCTGTGATGTGATGTGTCAGCAACATGAAGCAGTGTGTGGAATCCAAATGCCAT 2892
Db 2830 GGCAGCTGTGATGTGATGTGTCAGCAACATGAAGCAGTGTGTGGAATCCAAATGCCAT 2889
Qy 2893 GTGGCCTCTTCCCTTTGGCAGTGTATGGATGTTATAGATGAGACACCTGCCCT 2952
Db 2890 GTGGCCTCTTCCCTTTGGCAGTGTATGGATGTTATAGATGAGACACCTGCCCT 2949
Qy 2953 GAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTTGGAGCAACAGGCTGTGGCTGG 3012
Db 2950 GAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTTGGAGCAACAGGCTGTGGCTGG 3009
Qy 3013 TGTACTGATCCAGCAATCTGGGAAAGGAAATGATAGAGGTTCTTAAAGGACCA 3072
Db 3010 TGTACTGATCCAGCAATCTGGGAAAGGAAATGATAGAGGTTCTTAAAGGACCA 3069
Qy 3073 GTGAAGATGCTTCCAGACCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATTC 3132

Db 3070 GTGAAGATGCTTTCGCAAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTC 3129
Qy 3133 AGCATGTGTCTAGGAGACAGAGATCAAACTGTGTCTTTCAATTCACCTGTCCAGCTTCCAA 3192
Db 3130 AGCATGTGTCTAGGAGACAGAGATCAAACTGTGTCTTTCAATTCACCTGTCCAGCTTCCAA 3189
Qy 3193 TGCACGGCCACAGTAAATGCAATCAATCAGAGATCTGTGAGAACTGTGAGAACTGTGAGC 3252
Db 3190 TGCACGGCCACAGTAAATGCAATCAATCAGAGATCTGTGAGAACTGTGAGAACTGTGAGC 3249
Qy 3253 ACAGGCAAGCAGTTCGAGAGCCCTGCATATCTGGCTTCTACGGTATCCCACTAATGAGGG 3312
Db 3250 ACAGGCAAGCAGTTCGAGAGCCCTGCATATCTGGCTTCTACGGTATCCCACTAATGAGGG 3309
Qy 3313 AAATGTACGCCATGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 3372
Db 3310 AAATGTACGCCATGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 3369
Qy 3373 TGCCTCTGCACACCAAGGCGTCAAGGGGAGAGTGCAGCTATGTGAGGTAGAAAAT 3432
Db 3370 TGCCTCTGCACACCAAGGCGTCAAGGGGAGAGTGCAGCTATGTGAGGTAGAAAAT 3429
Qy 3433 CGATACCAAGGAAACCCCTCTCAGAGAAACATGTTATTAATTAATTAATTAATTAATTAATTAAT 3492
Db 3430 CGATACCAAGGAAACCCCTCTCAGAGAAACATGTTATTAATTAATTAATTAATTAATTAATTAAT 3489
Qy 3493 TTCACTTTAGTCTATCCAGGAGATGATCGTATTAACAGCTATCAATTTTGGCT 3552
Db 3490 TTCACTTTAGTCTATCCAGGAGATGATCGTATTAACAGCTATCAATTTTGGCT 3549
Qy 3553 ACTCTCTGACGAAACCAACAGGATTTGGACATGTTTCAATCAATGCTTCCAGAAATTTCAAC 3612
Db 3550 ACTCTCTGACGAAACCAACAGGATTTGGACATGTTTCAATCAATGCTTCCAGAAATTTCAAC 3609
Qy 3613 CTCAATCACTGGGCTGGCTGCTTCTCAGCTGGAACCCAGGCTGGAGAGATGCT 3672
Db 3610 CTCAATCACTGGGCTGGCTGCTTCTCAGCTGGAACCCAGGCTGGAGAGATGCT 3669
Qy 3673 GTTGTTCACAAACCAACATTAAGAGATACAAAGATAGTTTCTCTAATCAGAAATTTGAT 3732
Db 3670 GTTGTTCACAAACCAACATTAAGAGATACAAAGATAGTTTCTCTAATGAGAAATTTGAT 3729
Qy 3733 TTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTGAGTAATTTCACTTGGCCATC 3792
Db 3730 TTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTGAGTAATTTCACTTGGCCATC 3789
Qy 3793 ABAATTCAGATTCCTTCTCAGCAGACAGAAATTTTATGACCTGTGAGATTTCTTCGTCG 3852
Db 3790 ABAATTCAGATTCCTTCTCAGCAGACAGAAATTTTATGACCTGTGAGATTTCTTCGTCG 3849
Qy 3853 ACTTCTTCAGTTGTTTCTCTCTTTGCTCCTGGTGGCTGCTGTGTTTGGAGATCAAA 3912
Db 3850 ACTTCTTCAGTTGTTTCTCTCTTTGCTCCTGGTGGCTGCTGTGTTTGGAGATCAAA 3909
Qy 3913 CAAAGTTGTGGGCTCCAGACGTAGAGAGCAACTTTCTCGAGAGATGCAACAGATGGCC 3972
Db 3910 CAAAGTTGTGGGCTCCAGACGTAGAGAGCAACTTTCTCGAGAGATGCAACAGATGGCC 3969
Qy 3973 AGCCGCTCCCTTTCCTCTGTAAATGTGCGCTTGGAAACAGATGAGGAGCTCTCTGATCTT 4032
Db 3970 AGCCGCTCCCTTTCCTCTGTAAATGTGCGCTTGGAAACAGATGAGGAGCTCTCTGATCTT 4029
Qy 4033 ATTGGGGGAGATTAAGACTGTTCCTTCCAAACCCATTTGCACTGAGGAGCTGTTTGGCAAC 4092
Db 4030 ATTGGGGGAGATTAAGACTGTTCCTTCCAAACCCATTTGCACTGAGGAGCTGTTTGGCAAC 4089
Qy 4093 ABAAGCCGCTCTCTCTGTGTTTGTGAGCTCCCTCGAGGCTGGGTGGCATCCCTCT 4152
Db 4090 ABAAGCCGCTCTCTCTGTGTTTGTGAGCTCCCTCGAGGCTGGGTGGCATCCCTCT 4149
Qy 4153 CCTGGGAGTGTGCTGTGGCCAGCCCTGGTGACATTTCTCAGCAGATGCGG 4212
Db 4150 CCTGGGAGTGTGCTGTGGCCAGCCCTGGTGACATTTCTCAGCAGATGCGG 4209

QY 4213 ATAGTGTACAGGAGAGTACGAGCCGTGTGAAACCGGAAACGAGCAGCCCTGACAG 4272
DB 4210 ATAGTGTACAGGAGAGTACGAGCCGTGTGAAACCGGAAACGAGCAGCCCTGACAG 4269
QY 4273 CCTGGGACCTGCATCTGA 4290
DB 4270 CCTGGGACCTGCATCTGA 4287
RESULT 2
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-16
Query Match 76.6%; Score 3285; DB 3; Length 4072;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 313 TGTGTCAACGGCGGTGCTGCAACCGCTGCGACCGCGCAGTGCCTGCGCGCGCGGTGG 372
DB 310 TGTGTCAACGGCGGTGCTGCAACCGCTGCGACCGCGCAGTGCCTGCGCGCGCGGTGG 369
QY 373 GTGGCGGAGCAATGCCAGCACTGCGGGGCGCGCTTACAGTAACTGAGTCTTCTGGGTTT 432
DB 370 GTGGCGGAGCAATGCCAGCACTGCGGGGCGCGCTTACAGTAACTGAGTCTTCTGGGTTT 429
QY 433 GTGACAGATGGACCTGGAAATTTATAAATACAAACGAGTGCAGTGGCTCATTTGAAGGA 492
DB 430 GTGACAGATGGACCTGGAAATTTATAAATACAAACGAGTGCAGTGGCTCATTTGAAGGA 489
QY 493 CAGCCAAATAGAAATAGAGACTTGGTTCAATCATTTTGTCTACAGAGTGTAGTTGGGAC 552
DB 490 CAGCCAAATAGAAATAGAGACTTGGTTCAATCATTTTGTCTACAGAGTGTAGTTGGGAC 549
QY 553 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCAATTTAGTGGC 612
DB 550 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCAATTTAGTGGC 609
QY 613 CTCATTGTTCTGTAGAGAGATGGCAATGAGACTGTCTCCCTGAGGTTGTTGCCACATCAGGT 672
DB 610 CTCATTGTTCTGTAGAGAGATGGCAATGAGACTGTCTCCCTGAGGTTGTTGCCACATCAGGT 669
QY 673 TATGCCCTGTGCAATTTTATGATGCTGCTTATTAATTTGACTGGAATTAATTAAT 732
DB 670 TATGCCCTGTGCAATTTTATGATGCTGCTTATTAATTTGACTGGAATTAATTAAT 729
QY 733 TACAGTTTGTATGTGTCCAAATTAACCTGTCAGGCCGAGGAGAGTGTAAAGATCAGTAAT 792
DB 730 TACAGTTTGTATGTGTCCAAATTAACCTGTCAGGCCGAGGAGAGTGTAAAGATCAGTAAT 789
QY 793 AGCAGCGAAATCTGTTGAATGTGAATGTTCTGAAATCTGAAAGCGTGAAGCATGTGACATT 852
DB 790 AGCAGCGAATCTGTTGAATGTGAATGTTCTGAAATCTGAAAGCGTGAAGCATGTGACATT 849

QY 853 CCTCAGCTGTACAGCAACCTGTGGTTTTCTCTCATCGAGGCACTCTGCAATTCAGATGATGTC 912
DB 850 CCTCAGCTGTACAGCAACCTGTGGTTTTCTCTCATCGAGGCACTCTGCAATTCAGATGATGTC 909
QY 913 AGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTCGGATGTTAGTTCCTCTGACAGCT 972
DB 910 AGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTCGGATGTTAGTTCCTCTGACAGCT 969
QY 973 AACCACTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCAT 1032
DB 970 AACCACTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCAT 1029
QY 1033 AAAGCTGTGTCATATGGAACATTTATGCGGTGTTGGAGGATATATGTTCAACCACTCA 1092
DB 1030 AAAGCTGTGTCATATGGAACATTTATGCGGTGTTGGAGGATATATGTTCAACCACTCA 1089
QY 1093 GATTATAACATGGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCGT 1152
DB 1090 GATTATAACATGGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCGT 1149
QY 1153 TCTGTGAACAATGTGGTTGTTAGATATGTCATTTCTTTGGCAATTATACAGGATATAAAT 1212
DB 1150 TCTGTGAACAATGTGGTTGTTAGATATGTCATTTCTTTGGCAATTATACAGGATATAAAT 1209
QY 1213 TACATGTATGGAGGAAAAATTCATCCAACTGGGAATGTACCAATGAGTTGAGAGTTTTT 1272
DB 1210 TACATGTATGGAGGAAAAATTCATCCAACTGGGAATGTACCAATGAGTTGAGAGTTTTT 1269
QY 1273 CACATTCATAATGAGTCACTGGGTGTTTGAACCCCTAAGGCAAGAGAGCATGACATG 1332
DB 1270 CACATTCATAATGAGTCACTGGGTGTTTGAACCCCTAAGGCAAGAGAGCATGACATG 1329
QY 1333 GTTGGSCACTCTGCACACATTTACACTGAAGATGCGGAGTGGTCACTGCTGCTATC 1392
DB 1330 GTTGGSCACTCTGCACACATTTACACTGAAGATGCGGAGTGGTCACTGCTGCTATC 1389
QY 1393 TTTGGTCACTGCCCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG 1452
DB 1390 TTTGGTCACTGCCCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG 1449
QY 1453 AACACATGAGATATATTACACACCCAGGTCCTTGTGCAAGGGGTTACGCCATAGC 1512
DB 1450 AACACATGAGATATATTACACACCCAGGTCCTTGTGCAAGGGGTTACGCCATAGC 1509
QY 1513 AGTGTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTTGGTTCACAGGCTTTTCAGT 1572
DB 1510 AGTGTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTTGGTTCACAGGCTTTTCAGT 1569
QY 1573 GCCAATAAGTACCGGTTGACAGATCTCTACCGATATGATGTGGATACCCAGATGTGG 1632
DB 1570 GCCAATAAGTACCGGTTGACAGATCTCTACCGATATGATGTGGATACCCAGATGTGG 1629
QY 1633 ACCATTCTTAAAGACAGCCGATTTTTCGTTACTTTTGCACACAGCTGTGATAGTGGGA 1692
DB 1630 ACCATTCTTAAAGACAGCCGATTTTTCGTTACTTTTGCACACAGCTGTGATAGTGGGA 1689
QY 1693 ACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGGCCGCAAAA 1752
DB 1690 ACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGGCCGCAAAA 1749
QY 1753 TGCTTCTCTTCAGATTTTCATGCGCTTATGACATTTGTCAGCGCTGTGCTGCTTCCC 1812
DB 1750 TGCTTCTCTTCAGATTTTCATGCGCTTATGACATTTGTCAGCGCTGTGCTGCTTCCC 1809
QY 1813 AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTTGACAGTCTTTACAAACAGC 1872
DB 1810 AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTTGACAGTCTTTACAAACAGC 1869
QY 1873 ACCATGATGCTGTTGGTGGTTTCAATAGTCTCTCCTCAGGACATCTCTGTTATTCACC 1932
DB 1870 ACCATGATGCTGTTGGTGGTTTCAATAGTCTCTCCTCAGGACATCTCTGTTATTCACC 1929

QY	1933	TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTGTGTTTAGCAGCAGGACCTGGTATT	1992
Db	1930	TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTGTGTTTAGCAGCAGGACCTGGTATT	1989
QY	1993	CGGTGTGTGTGGAACACAGAGTGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAA	2052
Db	1990	CGGTGTGTGTGGAACACAGAGTGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAA	2049
QY	2053	CAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCAATGACAGATGT	2112
Db	2050	CAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCAATGACAGATGT	2109
QY	2113	GACCAGACACAGATGTTACAGCTGTACAGCTGTACAGCAACCAATGACTGCCACTGGTGAAT	2172
Db	2110	GACCAGACACAGATGTTACAGCTGTACAGCTGTACAGCAACCAATGACTGCCACTGGTGAAT	2169
QY	2173	GACCAATGTGTCTCCAGGAACCAAGCTGTCTAGAGGCGCAGATCTCCATTTTGTAGGTAT	2232
Db	2170	GACCAATGTGTCTCCAGGAACCAAGCTGTCTAGAGGCGCAGATCTCCATTTTGTAGGTAT	2229
QY	2233	GAGAAATTTGCCCCAGGATACCCCATGTACTGTATACAGAGACCAAGCTGCAGGAGC	2292
Db	2230	GAGAAATTTGCCCCAGGATACCCCATGTACTGTATACAGAGACCAAGCTGCAGGAGC	2289
QY	2293	TGTGCTCTGGACCAAGAACTGCACTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2352
Db	2290	TGTGCTCTGGACCAAGAACTGCACTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2349
QY	2353	GAAATATCTGTGCAATTTGGCTGGCATTTGGTTGGAACTCACTGTTTGAATTTACTACT	2412
Db	2350	GAAATATCTGTGCAATTTGGCTGGCATTTGGTTGGAACTCACTGTTTGAATTTACTACT	2409
QY	2413	GCCAAAGAGAAATTAACAAATGCTTAAATGTTCTGTAGGAACCAATGCGCTTTGGCT	2472
Db	2410	GCCAAAGAGAAATTAACAAATGCTTAAATGTTCTGTAGGAACCAATGCGCTTTGGCT	2469
QY	2473	TCTCTTCAACCCAGAGAGAGTGAATTTGCTCTTAAAGCAGCTGCGAATAATGCACTCA	2532
Db	2470	TCTCTTCAACCCAGAGAGAGTGAATTTGCTCTTAAAGCAGCTGCGAATAATGCACTCA	2529
QY	2533	TCTCAGACATGTCACCTTAACCCATGCGCTGGCTCGGAGATCAATGTG	2592
Db	2530	TCTCAGACATGTCACCTTAACCCATGCGCTGGCTCGGAGATCAATGTG	2589
QY	2593	TCCTACTGCTGTGGAGATATGTCCTTCAATATGTTTACTACAGTGGATCCG	2652
Db	2590	TCCTACTGCTGTGGAGATATGTCCTTCAATATGTTTACTACAGTGGATCCG	2649
QY	2653	TCTGAGCCAGTGTGATGCTGGAATTTGAGAAATTTATCAGAACCCAGTACTCGGGAGCTG	2712
Db	2650	TCTGAGCCAGTGTGATGCTGGAATTTTATCAGAACCCAGTACTCGGGAGCTG	2709
QY	2713	AAGGCTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGCCCTGCAAAACAC	2772
Db	2710	AAGGCTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGCCCTGCAAAACAC	2769
QY	2773	AGTGCTAAGCAGTGTCCGACACCACTGTGCTGTAGGACAGCATGTGGAGATGTCACAGC	2832
Db	2770	AGTGCTAAGCAGTGTCCGACACCACTGTGCTGTAGGACAGCATGTGGAGATGTCACAGC	2829
QY	2833	GSCAGCTGTGAGTGTATGCTGTGAGCAACATGAGCAGTGTGTGGATCTCCAAATGCTTAT	2892
Db	2830	GSCAGCTGTGAGTGTATGCTGTGAGCAACATGAGCAGTGTGTGGATCTCCAAATGCTTAT	2889
QY	2893	GTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACATGAGCACTGCCCGCT	2952
Db	2890	GTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACATGAGCACTGCCCGCT	2949
QY	2953	GAAATTTGTTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACCAAGCTGTGGCTGG	3012
Db	2950	GAAATTTGTTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACCAAGCTGTGGCTGG	3009
QY	3013	TGTACTGATCCACAGCAATACTGGCAAGGAAATGCAATAGAGGGTTCCTATAAAGGACCA	3072

Db	3010	TGTACTGATCCAGCAATACTGGCAAGGAAATGTCATAGAGGGTTCCTATAAAGGACCA	3069
QY	3073	GTGAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTC	3132
Db	3070	GTGAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTC	3129
QY	3133	AGCATGTGTCTAGAGCAGACAGATACAACTGGTCTTTTCAATCTGCTCCAGCTTGCCAA	3192
Db	3130	AGCATGTGTCTAGAGCAGACAGATACAACTGGTCTTTTCAATCTGCTCCAGCTTGCCAA	3189
QY	3193	TGCAACGGCCACAGTAAATGCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACC	3252
Db	3190	TGCAACGGCCACAGTAAATGCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACC	3249
QY	3253	ACAGCAAGCATCTGGAGACCTGCAATCTGCTGCTTCTAGGTCATCCCAATGGAGG	3312
Db	3250	ACAGCAAGCATCTGGAGACCTGCAATCTGCTGCTTCTAGGTCATCCCAATGGAGG	3309
QY	3313	AAATGTGACCCATGCAAGTGGCAATGGCAGCGCTCTCTGTGCAACCAACACGGGCAAG	3372
Db	3310	AAATGTGACCCATGCAAGTGGCAATGGCAGCGCTCTCTGTGCAACCAACACGGGCAAG	3369
QY	3373	TGCTTCTGACCAACCAAGGCGTCAAGGGGAGAGTCCAGCTATGTGAGCTAGAAAT	3432
Db	3370	TGCTTCTGACCAACCAAGGCGTCAAGGGGAGAGTCCAGCTATGTGAGCTAGAAAT	3429
QY	3433	CGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATCTCTTCTTATTCACATCAG	3492
Db	3430	CGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATCTCTTCTTATTCACATCAG	3489
QY	3493	TTACCTTTTGTCTATCCAGGAAGATGATGCTTATACAGCTATCAATTTTGTGGCT	3552
Db	3490	TTACCTTTTGTCTATCCAGGAAGATGATGCTTATACAGCTATCAATTTTGTGGCT	3549
QY	3553	ACTCTGACGAAACCAACAGGATTTGGACATGTTTCAATGCTCTCCAAAGAAATTTCAAC	3612
Db	3550	ACTCTGACGAAACCAACAGGATTTGGACATGTTTCAATGCTCTCCAAAGAAATTTCAAC	3609
QY	3613	CTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGAGAGATGCT	3672
Db	3610	CTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGAGAGATGCT	3669
QY	3673	GTGTTTTCAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATCAGAGTTTGTAT	3732
Db	3670	GTGTTTTCAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATCAGAGTTTGTAT	3729
QY	3733	TTTCGCAACCCCAATATCACTTTCTTTTGTGTTTATGTGAGTAAATTTACCTGGCCCATC	3792
Db	3730	TTTCGCAACCCCAATATCACTTTCTTTTGTGTTTATGTGAGTAAATTTACCTGGCCCATC	3789
QY	3793	AAAAATTCAG 3801	
Db	3790	AAAAATTCAG 3798	

RESULT 3

US-09-245-041-18
; Sequence 18, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 18

; LENGTH: 2625

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-245-041-18

Query Match

Best Local Similarity 42.8%; Score 1837; DB 3; Length 2625;

Matches 2037; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	313	TGTGTCAACGGGGTGGCTGCAACCGTGGCCAGCCGCGAGTGGCTGCGCCGCGCGGTGG	372
DB	310	TGTGTCAACGGGGTGGCTGCAACCGTGGCCAGCCGCGAGTGGCTGCGCCGCGCGGTGG	369
QY	373	GTGGCGAGCAATGCCAGCACTGGCGGGCGCCCTTCAGACTAACTGGATCTTCTGGGTTT	432
DB	370	GTGGCGAGCAATGCCAGCACTGGCGGGCGCCCTTCAGACTAACTGGATCTTCTGGGTTT	429
QY	433	GTGACAGATGGACCTGGAAATTAATAACAAACGAAAGTGACGCTGGCTCAITGAAGGA	492
DB	430	GTGACAGATGGACCTGGAAATTAATAACAAACGAAAGTGACGCTGGCTCAITGAAGGA	489
QY	493	CAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTGGAC	552
DB	490	CAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTGGAC	549
QY	553	CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC	612
DB	550	CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC	609
QY	613	CTCATGTTCTGAGAGATGGCAATCAGACTGTCCCTGAGGTGTGTCACATCAGGT	672
DB	610	CTCATGTTCTGAGAGATGGCAATCAGACTGTCCCTGAGGTGTGTCACATCAGGT	669
QY	673	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAATCT	732
DB	670	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAATCT	729
QY	733	TACAGTTTGTATGTGTCAAATAAATCTGCTCAGCCGAGAGAGTGTAGATCAGTAAT	792
DB	730	TACAGTTTGTATGTGTCAAATAAATCTGCTCAGCCGAGAGAGTGTAGATCAGTAAT	789
QY	793	AGCAGCAAACTGTTGAATGGAATGTTCTCAAACTGGAAGGTGAAGCATGTGACAT	852
DB	790	AGCAGCAAACTGTTGAATGGAATGTTCTCAAACTGGAAGGTGAAGCATGTGACAT	849
QY	853	CCTCACTGTACAGAACACTGTGGTTTCTCTCATCGAGGCATCTGCAATTTCAAGTGTGC	912
DB	850	CCTCACTGTACAGAACACTGTGGTTTCTCTCATCGAGGCATCTGCAATTTCAAGTGTGC	909
QY	913	AGAGATGCTCTGCTTCTCAGACTGGCAGGTCTCGATGTTTCAAGTGTGACAGCT	972
DB	910	AGAGATGCTCTGCTTCTCAGACTGGCAGGTCTCGATGTTTCAAGTGTGACAGCT	969
QY	973	AACAGTCAATTTGGACTCGAGAGAAATATTCTAACTTAAAGTCTCCCGAGAGCATCTCAT	1032
DB	970	AACAGTCAATTTGGACTCGAGAGAAATATTCTAACTTAAAGTCTCCCGAGAGCATCTCAT	1029
QY	1033	AAAGTGTGTTCAATGGAAACATTTATGTTGGTGTGTTGAGGATATATGTTCAACCACTCA	1092
DB	1030	AAAGTGTGTTCAATGGAAACATTTATGTTGGTGTGTTGAGGATATATGTTCAACCACTCA	1089
QY	1093	GATTATAACATGGTTCTAGCGTATGACCTTCTTAGGGAGTGGCTTCCACTAAGCCGT	1152
DB	1090	GATTATAACATGGTTCTAGCGTATGACCTTCTTAGGGAGTGGCTTCCACTAAGCCGT	1149
QY	1153	TCTGTGAACAAATGTTGTTTGTAGATATGTTGCTATTTTGGCATTTATCAAGGATAAAAT	1212
DB	1150	TCTGTGAACAAATGTTGTTTGTAGATATGTTGCTATTTTGGCATTTATCAAGGATAAAAT	1209
QY	1213	TACATGTATGAGGAAATTTGATCCAACTGGGAACTGTGACCAATGATGATGAGATTTT	1272

DB	1210	TACATGTATGAGGAAATTTGATTTCAACTGGGAATGTGACCAATGATGATGAGTGTGAGATTTT	1269
QY	1273	CACATTCAATAGATCATGGGTGTTGTTGACCCCTAAGGCAAGGAGCAGTATGCAATG	1332
DB	1270	CACATTCAATAGATCATGGGTGTTGTTGACCCCTAAGGCAAGGAGCAGTATGCAATG	1329
QY	1333	GTGGGCACTCTGCACACACATTTGTTACCTCAAGAAATGGCCGAGTGTGATGCTGCTCATC	1392
DB	1330	GTGGGCACTCTGCACACATTTGTTACCTCAAGAAATGGCCGAGTGTGATGCTGCTCATC	1389
QY	1393	TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGAAATATGATTTGGATTAAG	1452
DB	1390	TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGAAATATGATTTGGATTAAG	1449
QY	1453	AACACATGGAGTATATTACACACCCAGGGTGCCTTTGTGCAAGGGGTTACGGCCATAGC	1512
DB	1450	AACACATGGAGTATATTACACACCCAGGGTGCCTTTGTGCAAGGGGTTACGGCCATAGC	1509
QY	1513	AGTGTTTAGCAACATAGGACCAAGGCCCTATACGTTTCATGTTGGCTACAAAGCTTTTCAGT	1572
DB	1510	AGTGTTTAGCAACATAGGACCAAGGCCCTATACGTTTCATGTTGGCTACAAAGCTTTTCAGT	1569
QY	1573	GCCAAATAGTACCGCTTCAGATGATCTCTACCGATATGATGTCGATACCCAGATGTTGG	1632
DB	1570	GCCAAATAGTACCGCTTCAGATGATCTCTACCGATATGATGTCGATACCCAGATGTTGG	1629
QY	1633	ACCATTCTTTAAGGACAGCCGATTTTTCGTTTACTTTGCAACACAGCTGTGATAGTGGGA	1692
DB	1630	ACCATTCTTTAAGGACAGCCGATTTTTCGTTTACTTTGCAACACAGCTGTGATAGTGGGA	1689
QY	1693	ACCATGCTGTTGTTGGGGAAACACACACATGACATCTATGAGCCATGCGCCGAAA	1752
DB	1690	ACCATGCTGTTGTTGGGGAAACACACACATGACATCTATGAGCCATGCGCCGAAA	1749
QY	1753	TGCTTCTCTTCAGATTTTCATGGCTTATGACATTTGCTGACCGCTGTGATGCTGCTCC	1812
DB	1750	TGCTTCTCTTCAGATTTTCATGGCTTATGACATTTGCTGACCGCTGTGATGCTGCTCC	1809
QY	1813	AGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTTAGCAGTCTTTACACAAACG	1872
DB	1810	AGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTTAGCAGTCTTTACACAAACG	1869
QY	1873	ACATGATGCTGCTCGGTGTTTCAATAGTCTCTCCTCAGCGACATCTTGATTTTACC	1932
DB	1870	ACATGATGCTGCTCGGTGTTTCAATAGTCTCTCCTCAGCGACATCTTGATTTTACC	1929
QY	1933	TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTTTGTAGCAGCAGGACCTGGTAT	1992
DB	1930	TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTTTGTAGCAGCAGGACCTGGTAT	1989
QY	1993	CGGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGCACTGATGAA	2052
DB	1990	CGGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGCACTGATGAA	2049
QY	2053	CAAGAAGAAAGTTAAATCAGATTTTTCAAAAGAACTCTTGACCATGACAGATGT	2112
DB	2050	CAAGAAGAAAGTTAAATCAGATTTTTCAAAAGAACTCTTGACCATGACAGATGT	2109
QY	2113	GACGACACACAGATTTGTTACAGCTGTACAGCCACACCAATGACTGCACTGGTCAAT	2172
DB	2110	GACGACACACAGATTTGTTACAGCTGTACAGCCACACCAATGACTGCACTGGTCAAT	2169
QY	2173	GACCATTTGTGTCGCCGAAACACAGCTGCTCAGAAGGCCAGATCTCCATTTTGTAGTAT	2232
DB	2170	GACCATTTGTGTCGCCGAAACACAGCTGCTCAGAAGGCCAGATCTCCATTTTGTAGTAT	2229
QY	2233	GAGAAATGCCCCAAGGATAACCCCATGTAATCTGTTAAACAGAAAGACAGCTGGAGGC	2292
DB	2230	GAGAAATGCCCCAAGGATAACCCCATGTAATCTGTTAAACAGAAAGACAGCTGGAGGC	2289
QY	2293	TGTGCTCTGGACCAAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2352
DB	2290	TGTGCTCTGGACCAAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2349

QY	2353	G	2353
Db	2350	G	2350
RESULT 4			
US-09-245-041-12			
; Sequence 12, Application US/09245041			
; Patent No. 6274339			
; GENERAL INFORMATION:			
; APPLICANT: Moore, K.			
; APPLICANT: Nagle, D.			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT			
; FILE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY			
; FILE REFERENCE: 7853-136			
; CURRENT APPLICATION NUMBER: US/09/245,041			
; CURRENT FILING DATE: 1999-02-05			
; EARLIER APPLICATION NUMBER: 60/093,630			
; EARLIER FILING DATE: 1998-07-21			
; EARLIER APPLICATION NUMBER: 60/104,978			
; EARLIER FILING DATE: 1998-10-20			
; NUMBER OF SEQ ID NOS: 131			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 12			
; LENGTH: 6370			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-245-041-12			
Query Match			
Best Local Similarity 35.2%; Score 1510; DB 3; Length 6370;			
Matches 1510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	2781	GCAGTGC	CGGACACCATGTGCTTCAGGACAGCATGTGGAGATTGCACACGGCAGCTC 2840
Db	548	GCAGTGC	CGGACACCATGTGCTTCAGGACAGCATGTGGAGATTGCACACGGCAGCTC 607
QY	2841	TCAGTGC	ATGTGTCACACATCAAGCAGTGTGGACTCCATGCTATGTGGCCTC 2900
Db	608	TCAGTGC	ATGTGTCACACATCAAGCAGTGTGGACTCCATGCTATGTGGCCTC 667
QY	2901	CTTCCCTTTT	GGCCAGTGTATGGAATGTATACGATGACGACCTGCCCTCCCTGAAATTTG 2960
Db	668	CTTCCCTTTT	GGCCAGTGTATGGAATGTATACGATGACGACCTGCCCTCCCTGAAATTTG 727
QY	2961	TTCCAGCTACT	GTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTGCTACTGA 3020
Db	728	TTCCAGCTACT	GTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTGCTACTGA 787
QY	3021	TCCCAGCAAT	ACTGTCGCAAGGGAATGCATAGAGGTTCCCTATAAGGACAGTGAAGAT 3080
Db	788	TCCCAGCAAT	ACTGTCGCAAGGGAATGCATAGAGGTTCCCTATAAGGACAGTGAAGAT 847
QY	3081	GCCTTTCG	CAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTCAGCATGTG 3140
Db	848	GCCTTTCG	CAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTCAGCATGTG 907
QY	3141	TCTAGAG	CACAGATCAACTGTCTTTCATTCACGTCCAGCTTGCCATGCAACGG 3200
Db	908	TCTAGAG	CACAGATCAACTGTCTTTCATTCACGTCCAGCTTGCCATGCAACGG 967
QY	3201	CCACAGT	AAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAA 3260
Db	968	CCACAGT	AAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAA 1027
QY	3261	GCACCTCG	AGACTGCATATCTGGCTTCTACGGTATCCACCAATGGAGGAATGTCA 3320
Db	1028	GCACCTCG	AGACTGCATATCTGGCTTCTACGGTATCCACCAATGGAGGAATGTCA 1087
QY	3321	GCCATCA	AGTGCATAGGCGCTCTGTGCAACACCAACAGGCAAGTGTCTCTG 3380
Db	1088	GCCATCA	AGTGCATAGGCGCTCTGTGCAACACCAACAGGCAAGTGTCTCTG 1147

QY	2353	G 2353
Db	2350	G 2350
RESULT 4		
US-09-245-041-12		
; Sequence 12, Application US/09245041		
; Patent No. 6274339		
; GENERAL INFORMATION:		
; APPLICANT: Moore, K.		
; APPLICANT: Naale, D.		
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT		
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY		
; FILE REFERENCE: 7853-136		
; CURRENT APPLICATION NUMBER: US/09/245,041		
; CURRENT FILING DATE: 1999-02-05		
; EARLIER APPLICATION NUMBER: 60/093,630		
; EARLIER FILING DATE: 1998-07-21		
; EARLIER APPLICATION NUMBER: 60/104,978		
; EARLIER FILING DATE: 1998-10-20		
; NUMBER OF SEQ ID NOS: 131		
; SOFTWARE: Fast-SEQ for Windows Version 3.0		
; SEQ ID NO 12		
; LENGTH: 6370		
; TYPE: DNA		
; ORGANISM: Homo sapiens		
US-09-245-041-12		
Query Match 35.2%; Score 1510; DB 3; Length 6370;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	2781	GCAGTGGCGGACACCATGTGCTTTGAGGACAGCATGTGGAGATTGCACACGCGGACGCTC 2840
Db	548	GCAGTGGCGGACACCATGTGCTTTGAGGACAGCATGTGGAGATTGCACACGCGGACGCTC 607
QY	2841	TGAGTGCATGTGTGCGAGCAACATGAAGCAGTGTGTGGAATCCAAATGCCTATGTGGCTC 2900
Db	608	TGAGTGCATGTGTGCGAGCAACATGAAGCAGTGTGTGGAATCCAAATGCCTATGTGGCTC 667
QY	2901	CTTCCCTTTTGGCCAGTGTATGGAAATGGTATACCATGAGCAGCTGCCGCCCTGAAATTTG 2960
Db	668	CTTCCCTTTTGGCCAGTGTATGGAAATGGTATACCATGAGCAGCTGCCGCCCTGAAATTTG 727
QY	2961	TTCAGGCTACGTACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTGTACTGA 3020
Db	728	TTCAGGCTACGTACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTGTACTGA 787
QY	3021	TCCAGAGCAATACTGSCAAAAGGAAATGCATAGAGGGTTCTATATAAGGACCAAGTGAAGAT 3080
Db	788	TCCAGAGCAATACTGSCAAAAGGAAATGCATAGAGGGTTCTATATAAGGACCAAGTGAAGAT 847
QY	3081	GCCTTCGACAGCCCTACAGAAAATTTCTATCCACAGCCCTGTCTCAATTTCCAGCATGTG 3140
Db	848	GCCTTCGACAGCCCTACAGAAAATTTCTATCCACAGCCCTGTCTCAATTTCCAGCATGTG 907
QY	3141	TCTAGAGACAGCAGATACACTGTGCTTTTCATTCACCTGTCCAGCTTGGCAATGCAACGG 3200
Db	908	TCTAGAGACAGCAGATACACTGTGCTTTTCATTCACCTGTCCAGCTTGGCAATGCAACGG 967
QY	3201	CCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACTGTGACCAAGCAAA 3260
Db	968	CCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACTGTGACCAAGCAAA 1027
QY	3261	GCATGCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACCAATGAGGGGAATGTCA 3320
Db	1028	GCATGCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACCAATGAGGGGAATGTCA 1087
QY	3321	GCCATGCAAGTGCAATGGGCGAGCGTCTCTGTGCAACCAACCAACGCGGCAAGTCTTCTG 3380
Db	1088	GCCATGCAAGTGCAATGGGCGAGCGTCTCTGTGCAACCAACCAACGCGGCAAGTCTTCTG 1147

RESULT 5
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-1

Query Match 6.8%; Score 291; DB 3; Length 8827;
Best Local Similarity 100.0%; Pred. No. 2.1e-134;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4000 GCGTGGAAACAGATGAGGAGCTCTGATCTTATTGGGGGAGTATAAGACTGTTCCTCC 4059
DB 4266 GCGTGGAAACAGATGAGGAGCTCTGATCTTATTGGGGGAGTATAAGACTGTTCCTCC 4325

QY 4060 AAACCCATGCACTGGAGCCGTGTTTGGCAACAAAGCCGTGCTCTCTGTGTTTGTG 4119
DB 4326 AAACCCATGCACTGGAGCCGTGTTTGGCAACAAAGCCGTGCTCTCTGTGTTTGTG 4385

QY 4120 AGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCTGGGAGTCAAGTCTGTGTTGCC 4179
DB 4386 AGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCTGGGAGTCAAGTCTGTGTTGCC 4445

QY 4180 AGCGCCCTCGTGGACATTTCTCAGCAGATGCCGATGTTACAAAGGAGAAGTCAGGAGCC 4239
DB 4446 AGCGCCCTCGTGGACATTTCTCAGCAGATGCCGATGTTACAAAGGAGAAGTCAGGAGCC 4505

QY 4240 GTGAGAACCGGAAGCAGAGCCCTCGCAGAGCTGGAGCTGCATCTGA 4290
DB 4506 GTGAGAACCGGAAGCAGAGCCCTCGCAGAGCTGGAGCTGCATCTGA 4556

RESULT 6
US-09-702-705-966/c
; Sequence 966, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-966

Query Match 5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCACTCATCTCAGAGCATGTCACAGCTACCTTAACCCCATGGGTGCGC 2574
DB 246 CTGCGAATAATGCACTCATCTCAGAGCATGTCACAGCTACCTTAACCCCATGGGTGCGC 187

QY 2575 CTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTTACAAATAGT 2634
DB 186 CTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGTCTGAGCCCAAGTATGCTGGATTCGTGGAATTTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGTCTGAGCCCAAGTATGCTGGATTCGTGGAATTTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

RESULT 7
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCACTCATCTCAGAGCATGTCACAGCTACCTTAACCCCATGGGTGCGC 2574
DB 246 CTGCGAATAATGCACTCATCTCAGAGCATGTCACAGCTACCTTAACCCCATGGGTGCGC 187

QY 2575 CTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTTACAAATAGT 2634
DB 186 CTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGTCTGAGCCCAAGTATGCTGGATTCGTGGAATTTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGTCTGAGCCCAAGTATGCTGGATTCGTGGAATTTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

```
RESULT 8
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Query Match      5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 2574
DB 246 CTGCGAATAATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 2634
DB 186 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGCTGAGCCAGTGTGCTGGAAATTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGCTGAGCCAGTGTGCTGGAAATTTATCAGAA 67

QY 2695 CCAGTACTCGGGGACTGAAGGTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCAGTACTCGGGGACTGAAGGTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

RESULT 9
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966

Query Match      5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 2574
DB 246 CTGCGAATAATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 2634
DB 186 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGCTGAGCCAGTGTGCTGGAAATTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGCTGAGCCAGTGTGCTGGAAATTTATCAGAA 67

QY 2695 CCAGTACTCGGGGACTGAAGGTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCAGTACTCGGGGACTGAAGGTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

RESULT 9
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966

Query Match      5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 2574
DB 246 CTGCGAATAATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 2634
DB 186 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGCTGAGCCAGTGTGCTGGAAATTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGCTGAGCCAGTGTGCTGGAAATTTATCAGAA 67

QY 2695 CCAGTACTCGGGGACTGAAGGTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCAGTACTCGGGGACTGAAGGTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

RESULT 10
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

Query Match      1.3%; Score 55; DB 3; Length 90050;
Best Local Similarity 100.0%; Pred. No. 5.5e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3130 TCCAGCATGTCTTAGAGGACAGCAGATCAACTGGTCTTTCATTCACTGTCCAG 3184
DB 3184 TCCAGCATGTCTTAGAGGACAGCAGATCAACTGGTCTTTCATTCACTGTCCAG 31894

RESULT 11
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10

Query Match      1.3%; Score 55; DB 3; Length 90050;
Best Local Similarity 100.0%; Pred. No. 5.5e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3130 TCCAGCATGTCTTAGAGGACAGCAGATCAACTGGTCTTTCATTCACTGTCCAG 3184
DB 3184 TCCAGCATGTCTTAGAGGACAGCAGATCAACTGGTCTTTCATTCACTGTCCAG 31894
```


FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 1051
TYPE: DNA
ORGANISM: Mus musculus
US-09-245-041-10

Query Match 1.2%; Score 53; DB 3; Length 1051;
Best Local Similarity 100.0%; Pred. No. 4.4e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGG 419
DB 571 GCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGG 623

RESULT 12

US-09-245-041-8
Sequence 8, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 2419
TYPE: DNA
ORGANISM: Mus musculus
US-09-245-041-8

Query Match 1.1%; Score 47; DB 3; Length 2419;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1624 CAGATGTGACCATCTTAAAGCAGCGCGATTTTCCTTACTTGCA 1670
DB 1890 CAGATGTGACCATCTTAAAGCAGCGCGATTTTCCTTACTTGCA 1936

RESULT 13

US-09-245-041-4
Sequence 4, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 5973
TYPE: DNA
ORGANISM: Mus musculus
US-09-245-041-4

Query Match 0.8%; Score 35; DB 3; Length 5973;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 TGGGACCATTATATGTTTATGATGGGACTCAAT 581
DB 4379 TGGGACCATTATATGTTTATGATGGGACTCAAT 4413

RESULT 14

US-09-245-041-3/c
Sequence 3, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 17056
TYPE: DNA
ORGANISM: Mus musculus
US-09-245-041-3

Query Match 0.7%; Score 32; DB 3; Length 17056;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 TGTGACATTCCTCACTGTACAGCAACTGTGG 875
DB 3978 TGTGACATTCCTCACTGTACAGCAACTGTGG 3947

RESULT 15

US-09-245-041-23/c
Sequence 23, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 29
TYPE: DNA
ORGANISM: Artificial Sequence
US-09-245-041-23

Query Match 0.7%; Score 29; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 552 CCATTATATGTTTATGATGGGACTCAA 580
DB 29 CCATTATATGTTTATGATGGGACTCAA 1

RESULT 16
US-09-245-041-48
; Sequence 48, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Moore, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
US-09-245-041-48

Query Match 0.6%; Score 24; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1631 GGACCAATTCCTTAAGGACAGCCGAT 1654
DB 1 GGACCAATTCCTTAAGGACAGCCGAT 24

RESULT 17
US-09-319-989-1/c
; Sequence 1, Application US/09319989
; Patent No. 6190914
; GENERAL INFORMATION:
; APPLICANT: Grivell, Leslie A.
; APPLICANT: Teixeira De Mattos, Maarten J.
; APPLICANT: Blom, Jolanda
; TITLE OF INVENTION: METHODS FOR MODULATING METABOLIC PATHWAYS OF
; TITLE OF INVENTION: MICRO-ORGANISMS AND MICRO-ORGANISMS OBTAINABLE BY SAID
; TITLE OF INVENTION: METHODS
; FILE REFERENCE: 24615-20123.00
; CURRENT APPLICATION NUMBER: US/09/319,989
; CURRENT FILING DATE: 1999-06-14
; EARLIER APPLICATION NUMBER: PCT/NL97/00688
; EARLIER FILING DATE: 1997-12-12
; EARLIER APPLICATION NUMBER: EPO 96203520
; EARLIER FILING DATE: 1996-12-12
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1198
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA construct
; OTHER INFORMATION: encoding HAP2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(995)

US-09-319-989-1

Query Match 0.6%; Score 24; DB 3; Length 1198;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGTTGCT 203
DB 591 GCTGCTGCTGCTGCTGTTGCT 568

RESULT 18
US-09-086-663A-70/c
; Sequence 70, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA
; APPLICANT: KARSENTY, GERARD
; TITLE OF INVENTION: OSF2/CHAF1 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UTSC:525
; CURRENT APPLICATION NUMBER: US/09/086,663A
; CURRENT FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/080,189
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 60/048,430
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1644)
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-09-086-663A-70

Query Match 0.6%; Score 24; DB 4; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGTTGCT 203
DB 223 GCTGCTGCTGCTGCTGTTGCT 200

RESULT 19
US-08-997-685A-1/c
; Sequence 1, Application US/08997685A
; Patent No. 6551821
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University
; APPLICANT: Kandel, Eric
; TITLE OF INVENTION: Brain Cyclic Nucleotide Gated Ion Channel and Uses Thereof
; FILE REFERENCE: 0575/54806
; CURRENT APPLICATION NUMBER: US/08/997,685A
; CURRENT FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2733
; TYPE: DNA
; ORGANISM: mouse
US-08-997-685A-1

Query Match 0.6%; Score 24; DB 4; Length 2733;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGTTGCT 203

```
Db      2281 GCTGCTGCTGCTGCTGCTGCTGCTGCT 2258
; Sequence 1, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA
; APPLICANT: KARSENTY, GERARD
; TITLE OF INVENTION: OSP2/CBFA1 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UTSC:525
; CURRENT APPLICATION NUMBER: US/09/086,663A
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/080,189
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 60/048,430
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-09-086-663A-1

Query Match      0.6%; Score 24; DB 4; Length 3334;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203
Db 531 GCTGCTGCTGCTGCTGCTGCTGCTGCT 508

RESULT 21
US-09-245-041-24
; Sequence 24, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
US-09-245-041-24

Query Match      0.5%; Score 23; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 GCCAATAGATATGAGACTTC 517
Db 7 GCCAATAGATATGAGACTTC 29

RESULT 22
US-09-245-041-7/c

; Sequence 7, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-7

Query Match      0.5%; Score 23; DB 3; Length 221;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2713 AAGGTGCAACTGCATCAACCC 2735
Db 118 AAGGTGCAACTGCATCAACCC 96

RESULT 23
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-6

Query Match      0.5%; Score 23; DB 3; Length 226;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2713 AAGGTGCAACTGCATCAACCC 2735
Db 123 AAGGTGCAACTGCATCAACCC 101

RESULT 24
US-08-684-101-1/c
; Sequence 1, Application US/08684101
; Patent No. 5846769
; GENERAL INFORMATION:
; APPLICANT: Abrams, John M.
; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 2
```

```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: US  
ZIP: 77210  
  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/684.101  
FILING DATE: Concurrently herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UTXD:503/PAR  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577
```

Query Match 0.5%; Score 23; DB 2; Length 1677;
Best Local Similarity 100.0%; Pred.No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      180   GCTGCGTGTGGTGTTTCGC 202  
          |||  
Db       469   GCTGCGTGTGGTGTTTCGC 447
```

RESULT 25
US-09-205-814-1/c
Sequence 1, Application US/09205814
Patent No. 6211339
GENERAL INFORMATION:
APPLICANT: Abrams, John M.
TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE:
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/205,814
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US/08/684,101
APPLICATION NUMBER:

FILLING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTXD:503/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: na
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1677 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 313..727
US-09-205-814-1

Query Match 0.5%; Score 23; DB 3; Length 1677;
Best Local Similarity 100.0%; Pred.No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      180   GCTGCTGCTGCTGTTTGTGC 202  
          |||  
Db       469   GCTGCTGCTGCTGTTTGTGC 447
```

RESULT 26
US-09-245-041-30/c
Sequence 30, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
APPLICANT: Nagle, D.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
OF BODY WEIGHT DISORDERS INCLUDING OBESITY
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/093,630
EARLIER FILING DATE: 1998-07-21
EARLIER APPLICATION NUMBER: 60/104,978
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 26
TYPE: DNA
ORGANISM: Artificial Sequence
US-09-245-041-30

Query Match 0.5%; Score 22; DB 3; Length 26;
Best Local Similarity 100.0%; Pred.No. 0.98;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      854   CTCACGTGTACAGCAACTGTGG 875  
          |||  
Db       26   CTCACGTGTACAGCAACTGTGG 5
```

RESULT 27
US-09-621-976-7843
Sequence 7843, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.

Correspondence Address:
Addressee: Arnold, White & Durkee
Street: P.O. Box 4433
City: Houston
State: TX
Country: US
ZIP: 77210

Computer Readable Form:
Medium Type: Floppy disk
Computer: IBM PC compatible
Operating System: PC-DOS/MS-DOS
Software: Patent In Release #1.0, Version #1.30

Current Application Data:
Application Number: US/08/684,101
Filing Date: Concurrently herewith
Classification: 435

Attorney/Agent Information:
Name: Parker, David L.
Registration Number: 32,165
Reference/Docket Number: UTXD:503/PAR
Telecommunication Information:
Telephone: (512) 418-3000
Telefax: (512) 474-7577

Information for Seq ID No: 1:
Sequence Characteristics:
Length: 1677 base pairs
Type: nucleic acid
Strandedness: single
Topology: linear
Molecule Type: other nucleic acid
Description: /desc = "DNA"

Name/Key: CDS
Location: 313..727

US-08-684-101-1

Query Match 0.5%; Score 23; DB 2; Length 1677;
Best Local Similarity 100.0%; Pred.No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGC 202
Db 469 GCTGCTGCTGCTGCTGCTTTC 447

Result 25
US-09-205-814-1/c
Sequence 1, Application US/09205814
Patent No. 6211339

General Information:
Applicant: Abrams, John M.
Title of Invention: VERTEBRATE APOPTOSIS GENE:
Title of Invention: COMPOSITIONS AND METHODS
Number of Sequences: 2

Correspondence Address:
Addressee: Arnold, White & Durkee
Street: P.O. Box 4433
City: Houston
State: TX
Country: US
ZIP: 77210

Computer Readable Form:
Medium Type: Floppy disk
Computer: IBM PC compatible
Operating System: PC-DOS/MS-DOS
Software: Patent In Release #1.0, Version #1.30

Current Application Data:
Application Number: US/09/205,814
Filing Date:
Classification:
Prior Application Number: US/08/684,101

Filing Date:
Attorney/Agent Information:
Name: Parker, David L.
Registration Number: 32,165
Reference/Docket Number: UTXD:503/PAR
Telecommunication Information:
Telephone: (512) 418-3000
Telefax: (512) 474-7577

Information for Seq ID No: 1:
Sequence Characteristics:
Length: 1677 base pairs
Type: nucleic acid
Strandedness: single
Topology: linear
Molecule Type: other nucleic acid
Description: /desc = "DNA"

Name/Key: CDS
Location: 313..727

US-09-205-814-1/c

Query Match 0.5%; Score 23; DB 2; Length 1677;
Best Local Similarity 100.0%; Pred.No. 0.38;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGC 202
Db 469 GCTGCTGCTGCTGCTGCTTTC 447

Result 26
US-09-245-041-30/c
Sequence 30, Application US/09245041
Patent No. 6274339

General Information:
Applicant: Moore, K.
Title of Invention: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF OBESITY
Title of Invention: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
File Reference: 7853-136
Current Application Number: US/09/245,041
Current Filing Date: 1999-02-05
Earlier Application Number: 60/093,630
Earlier Filing Date: 1998-07-21
Earlier Application Number: 60/104,978
Earlier Filing Date: 1998-10-20
Number of Seq ID NOS: 131
Software: FastSeq for Windows Version 3.0
Seq ID NO 30
Length: 26
Type: DNA
Organism: Artificial Sequence

US-09-245-041-30

Query Match 0.5%; Score 22; DB 3; Length 26;
Best Local Similarity 100.0%; Pred.No. 0.98;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CTCACGTGTACAGCAACTGTGG 875
Db 26 CTCACGTGTACAGCAACTGTGG 5

Result 27
US-09-621-976-7843
Sequence 7843, Application US/09621976
Patent No. 6639063

General Information:
Applicant: Dumas Milne Edwards, J.B.
Applicant: Jobert, S.
Applicant: Giordano, J.Y.
Title of Invention: ESTs and Encoded Human Proteins.

```
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7843
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-7843

Query Match      0.5%; Score 22; DB 4; Length 365;
Best Local Similarity 100.0%; Pred. NO. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 GGCGGGGGGGGGGGGGGGGG 267
    |||||
Db 119 GGCGGGGGGGGGGGGGGG 140

RESULT 28
US-09-148-545-36
; Sequence 36, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
```

[illegible]

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 459
Query Match 0.5%; Score 22; DB 4; Length 459;
Best Local Similarity 100.0%; Pred.No.1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0;

QY 211 CCGCTGTGCTGCTGCTGCTGC 232
Db 34 CCGCTGTGCTGCTGCTGCTGC 55

RESULT 29
US-08-998-416-915/c
; Sequence 915, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgен
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 915:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG:570RP
; US-08-998-416-915
Query Match 0.5%; Score 22; DB 3; Length 688;
Best Local Similarity 100.0%; Pred.No.1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0;

QY 180 GCTGCTGCTGCTGCTGTTGTTG 201
Db 652 GCTGCTGCTGCTGCTGTTGTTG 631

```

```

RESULT 30
US-08-195-152-1/c
; Sequence 1, Application US/08195152
; Patent No. 5679541
; GENERAL INFORMATION:
; APPLICANT: Bonini, Nancy M.
; APPLICANT: Leiserson, William M.
; APPLICANT: Benzer, Seymour
; TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,152
; FILING DATE: 14-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59551/RPT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; US-08-195-152-1

```

```

Query Match 0.5%; Score 22; DB 1; Length 3231;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 182 TCCTGCTGCTGCTGTTGTGCT 203
Db 580 TCCTGCTGCTGCTGTTGTGCT 559

```

```

RESULT 31
US-08-754-311B-1/c
; Sequence 1, Application US/08754311B
; Patent No. 6623937
; GENERAL INFORMATION:
; APPLICANT: Bonini, Nancy M.
; APPLICANT: Leiserson, William M.
; APPLICANT: Benzer, Seymour
; TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/754,311B
; FILING DATE: 21-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/195,152
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59551/RPT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; US-08-754-311B-1

```

```

Query Match 0.5%; Score 22; DB 4; Length 3231;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 182 TCCTGCTGCTGCTGTTGTGCT 203
Db 580 TCCTGCTGCTGCTGTTGTGCT 559

```

```

RESULT 32
US-08-785-310A-3
; Sequence 3, Application US/08785310A
; Patent No. 5840532
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; TITLE OF INVENTION: Neuronal PAS Domain Protein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,310A
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: OSWAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4010 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```



```
; MOLECULE TYPE: cdna
US-08-785-310A-3

Query Match      0.5%; Score 22; DB 2; Length 4010;
Best Local Similarity 100.0%; Pred. No. 1.2; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 247 GCCGCGCGCGCGCGCGCGCGCG 268
DB 94 GCCGCGCGCGCGCGCGCGCGCG 115

RESULT 33
US-09-491-356C-7/c
; Sequence 7, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Glims, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491.356C
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 6558
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-491-356C-7

Query Match      0.5%; Score 22; DB 4; Length 6558;
Best Local Similarity 100.0%; Pred. No. 1.3; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 180 GCTGCTGCTGCTGCTGTTG 201
DB 6184 GCTGCTGCTGCTGCTGTTG 6163

RESULT 34
US-09-043-303-13/c
; Sequence 13, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUJII, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Alaxia Type 2 and
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043.303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-043-303-13

Query Match      0.5%; Score 21; DB 3; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 178 CGGCTGCTGCTGCTGCTGTTG 198
|||||
```

```
DB 60 CGGCTGCTGCTGCTGCTGTTG 40

RESULT 35
US-09-621-976-8302/c
; Sequence 8302, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8302
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8302

Query Match      0.5%; Score 21; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 3.3; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 248 CCGCGCGCGCGCGCGCGCGG 268
|||||
DB 36 CCGCGCGCGCGCGCGCGCGG 16

RESULT 36
US-09-621-976-12134
; Sequence 12134, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 12134
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-12134

Query Match      0.5%; Score 21; DB 4; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 250 GCGGCGCGCGCGCGCGCGGTG 270
DB 9 GCGGCGCGCGCGCGCGCGGTG 29

RESULT 37
US-09-621-976-17545/c
; Sequence 17545, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SEQ ID NO 17545
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17545

Query Match      0.5%; Score 21; DB 4; Length 257;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGCGGTGCGG 275
      |||||
Db 96 GCGCGCGCGCGGTGCGG 76

RESULT 38
US-09-370-838-145
; Sequence 145, Application US/09370838
; Patent No 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 145
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapien.
US-09-370-838-145

Query Match      0.5%; Score 21; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGCGCGCGCGCGCGGTG 270
      |||||
Db 57 GCGCGCGCGCGCGCGGTG 77

RESULT 39
US-08-234-783-1/c
; Sequence 1, Application US/08234783
; Patent No. 5622835
; GENERAL INFORMATION:
; APPLICANT: Herlyn, Meenhard
; APPLICANT: Morris, Jennifer
; APPLICANT: Rauscher III, Frank J.
; APPLICANT: Rodeck, Ulrich
; TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of
; TITLE OF INVENTION: Use Therefor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,783
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,783
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST48AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; US-08-234-783-1

Query Match      0.5%; Score 21; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGCGCGCGCGCGCGGTG 270
      |||||
Db 229 GCGCGCGCGCGCGCGGTG 209

RESULT 40
US-08-456-907-1/c
; Sequence 1, Application US/08456907
; Patent No. 5633142
; GENERAL INFORMATION:
; APPLICANT: Herlyn, Meenhard
; APPLICANT: Morris, Jennifer
; APPLICANT: Rauscher III, Frank J.
; APPLICANT: Rodeck, Ulrich
; TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of
; TITLE OF INVENTION: Use Therefor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/456,907
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,783
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST48AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 633 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..630
US-08-456-907-1

Query Match 0.5%; Score 21; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTG 270
Db 229 GCGGCGGCGGCGGCGGCGGTG 209

RESULT 41

PCT-US95-05523-1/c
; Sequence 1, Application PC/TUS9505523
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of Anatomy and Biology
; TITLE OF INVENTION: Wtl Monoclonal Antibodies and
; TITLE OF INVENTION: Methods of Use Therefor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: PCT/US95/05523
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,783
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REFERENCE/DOCKET NUMBER: WST48PCT
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
PCT-US95-05523-1

Query Match 0.5%; Score 21; DB 5; Length 633;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTG 270
Db 229 GCGGCGGCGGCGGCGGCGGTG 209

RESULT 42

US-08-835-099A-16
; Sequence 16, Application US/08835099A
; Patent No. 5874277
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,099A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 083649/1996
; FILING DATE: 05-APR-1996
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 640 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-835-099A-16

Query Match 0.5%; Score 21; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTG 270
Db 125 GCGGCGGCGGCGGCGGCGGTG 145

RESULT 43

US-09-157-349-16
; Sequence 16, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street

```
; CITY: Boston
; STATE: MA USA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-835-099A-10
;
; Query Match 0.5%; Score 21; DB 3; Length 640;
; Best Local Similarity 100.0%; Pred. No. 3.6;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 250 GCGGCGGCGGCGGCGGCGGTG 270
; DB 125 GCGGCGGCGGCGGCGGCGGTG 145
;
; RESULT 44
; US-08-835-099A-10
; Sequence 10, Application US/08835099A
; Patent No. 5874277
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,099A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 083649/1996
;
; CITY: Boston
; STATE: MA USA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-157-349-16
;
; Query Match 0.5%; Score 21; DB 2; Length 804;
; Best Local Similarity 100.0%; Pred. No. 3.6;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 250 GCGGCGGCGGCGGCGGCGGTG 270
; DB 125 GCGGCGGCGGCGGCGGCGGTG 145
;
; RESULT 45
; US-09-157-349-10
; Sequence 10, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-157-349-10

Query Match          0.5%; Score 21; DB 3; Length 804;
Best Local Similarity 100.0%; Pred.No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGGCGGCGGCGGCGGTG 270
      |||||
Db 125 GCGGCGGCGGCGGCGGCGGTG 145

RESULT 46
US-08-118-200-1
; Sequence 1, Application US/08118200
; Patent No. 6197500
; GENERAL INFORMATION:
; APPLICANT: SUTHERLAND, Grant R
; APPLICANT: RICHARDS, Robert I
; APPLICANT: SCHLESSINGER, David
; APPLICANT: NAGARAJA, Ramiah
; APPLICANT: KREMER, Eric J
; APPLICANT: YU, Sui
; APPLICANT: BAKER, Elizabeth
; APPLICANT: MULLEY, John C
; APPLICANT: MANDEL, Jean-Louis
; APPLICANT: PRITCHARD, Melanie April
; APPLICANT: LYNCH, Michael
; TITLE OF INVENTION: DNA SEQUENCES RELATED TO ISOLATED
; TITLE OF INVENTION: FRAGILE X SYNDROME
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,200
; FILING DATE: 09-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/802,650
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,232
; FILING DATE: 20-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/638,518
; FILING DATE: 04-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/966,517
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 020160-164
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-118-200-1

Query Match          0.5%; Score 21; DB 3; Length 1028;
Best Local Similarity 100.0%; Pred.No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGGCGGCGGCGGCGGTG 270
      |||||
Db 390 GCGGCGGCGGCGGCGGCGGTG 410

RESULT 47
US-08-458-745-1
; Sequence 1, Application US/08458745
; Patent No. 6242576
; GENERAL INFORMATION:
; APPLICANT: SUTHERLAND, Grant R
; APPLICANT: RICHARDS, Robert I
; APPLICANT: SCHLESSINGER, David
; APPLICANT: NAGARAJA, Ramiah
; APPLICANT: KREMER, Eric J
; APPLICANT: YU, Sui
; APPLICANT: BAKER, Elizabeth
; APPLICANT: MULLEY, John C
; APPLICANT: MANDEL, Jean-Louis
; APPLICANT: PRITCHARD, Melanie April
; APPLICANT: LYNCH, Michael
; TITLE OF INVENTION: DNA SEQUENCES RELATED TO ISOLATED
; TITLE OF INVENTION: FRAGILE X SYNDROME
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,745
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/118,200
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: US 07/802,650
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,232
; FILING DATE: 20-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/638,518
; FILING DATE: 04-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/966,517
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 020160-164
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1028 base pairs
; TYPE: nucleic acid
```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-458-745-1

Query Match          0.5%; Score 21; DB 3; Length 1028;
Best Local Similarity 100.0%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 250 GCGGCGGCGGCGGCGGCGGTG 270
DB 390 GCGGCGGCGGCGGCGGCGGTG 410

RESULT 48
US-08-765-875-1
; Sequence 1, Application US/08765875
; Patent No. 5914234
; GENERAL INFORMATION:
; APPLICANT: LEE, SE-JIN
; APPLICANT: MCPHERSON, ALEXANDRA C.
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,875
; FILING DATE:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,958
; FILING DATE:
; FILING DATE:
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3641
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HUMAN GDF-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 54..1274
US-08-765-875-1

Query Match          0.5%; Score 21; DB 2; Length 1393;
Best Local Similarity 100.0%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 248 CCGCGCGGCGGCGGCGGCGG 268
DB 136 CCGCGCGGCGGCGGCGGCGG 156

RESULT 49
US-08-795-671-1
; Sequence 1, Application US/08795671
; Patent No. 6008434
; GENERAL INFORMATION:
; APPLICANT: Se-Jin Lee and Alexandra McPherron
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,671
; FILING DATE: February 6, 1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/106001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HUMAN GDF-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 54..1274
US-08-795-671-1

Query Match          0.5%; Score 21; DB 3; Length 1393;
Best Local Similarity 100.0%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 248 CCGCGCGGCGGCGGCGGCGG 268
DB 136 CCGCGCGGCGGCGGCGGCGG 156

RESULT 50
US-09-454-540-1
; Sequence 1, Application US/09454540
; Patent No. 6517835
; GENERAL INFORMATION:
; APPLICANT: Se-Jin Lee and Alexandra McPherron
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/454,540
FILING DATE: 06-DEC-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,671
FILING DATE: February 6, 1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, PH.D., LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/106001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1393 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HUMAN GDF-11
FEATURE:
NAME/KEY: CDS
LOCATION: 54..1274
US-09-454-540-1

Query Match 0.5%; Score 21; DB 4; Length 1393;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGG 268
DB 136 CCGCGCGCGCGCGCGG 156

RESULT 51
US-09-626-896-24
Sequence 24, Application US/09628896
Patent No. 6656475
GENERAL INFORMATION:
APPLICANT: Lee, Se-Jin
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR RECEPTORS.
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS THEREOF, AND METHODS OF USING SAME
FILE REFERENCE: JHU1470-2
CURRENT APPLICATION NUMBER: US/09/626,896
CURRENT FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: 09/485,046
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: PCT/US98/15598
PRIOR FILING DATE: 1998-07-28
PRIOR APPLICATION NUMBER: 60/054,461
PRIOR FILING DATE: 1997-08-01
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 1393
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (54)...(1274)
OTHER INFORMATION: GDF-11
US-09-626-896-24

Query Match 0.5%; Score 21; DB 4; Length 1393;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGG 268
DB 136 CCGCGCGCGCGCGCGG 156
RESULT 52
US-09-220-132-82
Sequence 82, Application US/09220132
Patent No. 6506607
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 82
LENGTH: 1682
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-82

Query Match 0.5%; Score 21; DB 4; Length 1682;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGCGCGCGCGCGCGG 270
DB 746 GCGCGCGCGCGCGG 766

RESULT 53
US-08-857-636-58
Sequence 58, Application US/08857636
Patent No. 6552181
GENERAL INFORMATION:
APPLICANT: Dean, Michael Carlton
APPLICANT: Hahn, Heidi Eve
APPLICANT: Wicking, Carol
APPLICANT: Christiansen, Jeffrey G.
APPLICANT: Zaphitopoulos, Peter G.
APPLICANT: Gallani, Mae R.
APPLICANT: Shanley, Susan Mary
APPLICANT: Chidambaram, Abirami
APPLICANT: Vorechovsky, Igor
APPLICANT: Holmberg-Lindstrom, Erika
APPLICANT: Unden, Anne Birgitte
APPLICANT: Gillies, Susan Alana
APPLICANT: Negus, Kylie
APPLICANT: Smyth, Ian Mcleod
APPLICANT: Pressman, Carol Leah
APPLICANT: Lefell, David J.
APPLICANT: Gerrard, Bernard
APPLICANT: Goldstein, Alisa Miriam
APPLICANT: Mainwright, Brandon
APPLICANT: Toftgard, Rune Carl-Magnus
APPLICANT: Chenevix-Trench, Georgia
APPLICANT: Bale, Allen E.
TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1860
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-11

Query Match 0.5%; Score 21; DB 4; Length 1860;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGGTGTCGGG 275
DB 93 GCGCGCGCGCGGTGTCGGG 73

RESULT 56

US-09-220-132-67/c
Sequence 67, Application US/09220132
Patent No. 6506607
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 67
LENGTH: 1860
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-67

Query Match 0.5%; Score 21; DB 4; Length 1860;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGGTGTCGGG 275
DB 93 GCGCGCGCGCGGTGTCGGG 73

RESULT 57

US-09-220-132-137/c
Sequence 137, Application US/09220132
Patent No. 6506607
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
FILE REFERENCE: 07334-074001
CURRENT APPLICATION NUMBER: US/09/220,132
CURRENT FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 137
LENGTH: 1860
TYPE: DNA
ORGANISM: Homo sapiens
US-09-220-132-137

Query Match 0.5%; Score 21; DB 4; Length 1860;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGGTGTCGGG 275
DB 93 GCGCGCGCGCGGTGTCGGG 73

RESULT 58

US-08-937-067-5
Sequence 5, Application US/08937067
Patent No. 6433155
GENERAL INFORMATION:
APPLICANT: Umansky, Samuel
TITLE OF INVENTION: A FAMILY OF GENES ENCODING
TITLE OF INVENTION: APOPTOSIS-RELATED PEPTIDES; PEPTIDES ENCODED THEREBY AND
TITLE OF INVENTION: METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,067
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 23647-20018.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 216..1166
US-08-937-067-5

Query Match 0.5%; Score 21; DB 4; Length 1984;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGGTGTCGGG 268
DB 94 CCGCGCGCGCGGTGTCGGG 114

```
RESULT 59
US-07-952-800-1
; Sequence 1, Application US/07952800
; Patent No. 5403925
; GENERAL INFORMATION:
; APPLICANT: OZATO, KEIKO
; TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR HORMONE
; TITLE OF INVENTION: RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND
; STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/952,800
; FILING DATE: 19920928
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-21-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2130 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: DAUDI
; CELL TYPE: T cell lymphoma
; CELL LINE: DAUDI
; IMMEDIATE SOURCE:
; LIBRARY: cDNA DAUDI
; CLONE: DAUDI6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 180..1778
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 2109..2121
US-07-952-800-1
Query Match 0.5%; Score 21; DB 1; Length 2130;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 250 GCGGCGGCGGCGGCGCGGTG 270
Db 318 GCGGCGGCGGCGGCGCGGTG 338
RESULT 60
US-09-220-132-80/c
; Sequence 80, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
```

```
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 2634
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-80
Query Match 0.5%; Score 21; DB 4; Length 2634;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 248 CCGCGCGCGCGCGCGCGG 268
Db 25 CCGCGCGCGCGCGCGCGG 5
RESULT 61
US-09-620-312D-713
; Sequence 713, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 713
; LENGTH: 3169
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)..(2146)
US-09-620-312D-713
Query Match 0.5%; Score 21; DB 4; Length 3169;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 248 CCGCGCGCGCGCGCGCGG 268
Db 544 CCGCGCGCGCGCGCGCGG 564
```

```
RESULT 62
US-09-799-875-6/c
; Sequence 6, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-799-875-6
Query Match 0.5%; Score 21; DB 4; Length 3609;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 GCTGTGCTGCTGCTGCTGCC 233
Db 2674 GCTGTGCTGCTGCTGCTGCC 2654

RESULT 63
US-09-620-312D-348/c
; Sequence 348, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784C1P2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PC_FL_genes Version 1.0
; SEQ ID NO 348
; LENGTH: 4062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (25)..(2976)
US-09-620-312D-348
Query Match 0.5%; Score 21; DB 4; Length 4062;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 CGGCTGCTGCTGCTGCTGTTG 198
Db 2892 CGGCTGCTGCTGCTGCTGTTG 2872

RESULT 64
US-08-540-406-18
; Sequence 18, Application US/08540406
; Patent No. 5837538
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-540-406-18
Query Match 0.5%; Score 21; DB 2; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 CCGCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGCGG 373

RESULT 65
US-08-656-055-18
; Sequence 18, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/540,406
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-656-055-18

Query Match 0.5%; Score 21; DB 3; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGCGCGG 373

RESULT 66
US-08-954-668-18
Sequence 18, Application US/08954668
Patent No. 6172200
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley, Hoag & Eliot
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,668
FILING DATE: 20-Oct-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P
REGISTRATION NUMBER: 36709
```

```

REFERENCE/DOCKET NUMBER: SUV-003.06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-954-668-18

Query Match 0.5%; Score 21; DB 3; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGCGCGG 373

RESULT 67
US-08-918-658-18
Sequence 18, Application US/08918658
Patent No. 6429354
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,658
FILING DATE: 22-Aug-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/656,055
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/540,406
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-918-658-18

Query Match 0.5%; Score 21; DB 4; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 248 CCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGG 373

RESULT 68
US-09-724-631-18
; Sequence 18, Application US/09724631
; Patent No. 6551782
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; GOODRICH, LISA V
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/724,631
; FILING DATE: 28-Nov-85
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE: 1996-05-31
; APPLICATION NUMBER: 08/540,406
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1389
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-724-631-18

Query Match 0.5%; Score 21; DB 4; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 CCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGG 373

RESULT 69
US-08-954-701A-18
; Sequence 18, Application US/08954701A
; Patent No. 6610507
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; GOODRICH, LISA V
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-954-701A-18

Query Match 0.5%; Score 21; DB 4; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 CCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGG 373

RESULT 70
PCT-US95-13233-18
; Sequence 18, Application PC/TUS9513233
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13233
; FILING DATE: 06-OCT-1990
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1389
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US95-13233-18

Query Match
Best Local Similarity 100.0%; Pred. No. 3.9; Length 5288;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGCG 268
Db 353 CCGCGCGCGCGCGCGCGCGCG 373

RESULT 71
US-09-799-875-4/c
; Sequence 4, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/655,287
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 5983
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(3723)
US-09-799-875-4

Query Match
Best Local Similarity 100.0%; Pred. No. 3.9; Length 5983;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 GCTGTTGCTGCTGCTGCTGCC 233
Db 2788 GCTGTTGCTGCTGCTGCTGCC 2768

RESULT 72
US-08-457-273B-41/c
; Sequence 41, Application US/08457273B
; Patent No. 5849995
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael
; APPLICANT: Lin, Biaoyang
; APPLICANT: Nasir, Jamal
; TITLE OF INVENTION: Mouse Model for Huntington's Disease and
; TITLE OF INVENTION: Related DNA Sequences
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 5849995th Carolina
; COUNTRY: US
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,273B
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 3477-85A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10348 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-457-273B-41

Query Match
Best Local Similarity 100.0%; Pred. No. 4; Length 10348;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGCGCGCGCGCGCGTG 270
Db 460 GCGGCGCGCGCGCGCGCGTG 440

RESULT 73
US-08-556-419-13/c
; Sequence 13, Application US/08556419C
; Patent No. 6093549
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107.52271
; CURRENT APPLICATION NUMBER: US/08/556,419C
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 10348
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-556-419-13

Query Match
Best Local Similarity 100.0%; Pred. No. 4; Length 10348;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGCGCGCGCGCGCGTG 270
Db 460 GCGGCGCGCGCGCGCGCGTG 440

RESULT 74
US-09-041-886-14/c
; Sequence 14, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bradesen, Dale E.
; APPLICANT: Rabilzadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
```


TITLE OF INVENTION: Polypeptides and Methods of Use
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 10348 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 316..9748
US-09-041-886-14

Query Match 0.5%; Score 21; DB 3; Length 10348;
Best Local Similarity 100.0%; Pred.No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGGCGGGCGGGCGGGCGGTG 270
|||||
DB 460 GCGGGCGGGCGGGCGGGCGGTG 440

RESULT 75
US-08-246-982A-5/c
Sequence 5, Application US/08246982A
Patent No. 5686288
GENERAL INFORMATION:
APPLICANT: MacDonald, Marcy E.
APPLICANT: Ambrose, Christine M.
APPLICANT: Duyao, Mabel P.
APPLICANT: Gusella, James F.
TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,982A
FILING DATE: May 20, 1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge, A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0609.3880002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10366 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 316..9748
US-08-246-982A-5

Query Match 0.5%; Score 21; DB 1; Length 10366;
Best Local Similarity 100.0%; Pred.No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGGCGGGCGGGCGGGCGGTG 270
|||||
DB 460 GCGGGCGGGCGGGCGGGCGGTG 440

Search completed: March 1, 2004, 20:50:42
Job time : 328 secs

This Page Blank (uspto)

QY 1441 GATTGGATAAGAACACATGGAGTATATTACACACCAGGGTGCCTTGTGCAAGGGGT 1500
Db 1441 GATTGGATAAGAACACATGGAGTATATTACACACCAGGGTGCCTTGTGCAAGGGGT 1500
QY 1501 TACGGCCATAGCAGTGTTCAGACCATAGACACAGGAGCCCTATACGTTTCATGTGGCTAC 1560
Db 1501 TACGGCCATAGCAGTGTTCAGACCATAGGACACAGGAGCCCTATACGTTTCATGTGGCTAC 1560
QY 1561 AAGGCTTTTCAGTGCACCAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGAT 1620
Db 1561 AAGGCTTTTCAGTGCACCAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGAT 1620
QY 1621 ACCCAGATGTGGACCATCTTAAAGACAGCCGATTTTCGGTACTTTCAGACACAGCTGTG 1680
Db 1621 ACCCAGATGTGGACCATCTTAAAGACAGCCGATTTTCGGTACTTTCAGACACAGCTGTG 1680
QY 1681 ATAGTGAAGTGAACCATCTCGTGTGGGGGAAACACACACATATGATCATGAGC 1740
Db 1681 ATAGTGAAGTGAACCATCTCGTGTGGGGGAAACACACACATATGATCATGAGC 1740
QY 1741 CATGGCGCAAAATGCTTCTTTCAGATTTTCATGCGCTTANGCATGTGCTGAGCCGTGG 1800
Db 1741 CATGGCGCAAAATGCTTCTTTCAGATTTTCATGCGCTTANGCATGTGCTGAGCCGTGG 1800
QY 1801 TCAGTGTCTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Db 1801 TCAGTGTCTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
QY 1861 TTACACAAACAGCACCATGATGTGTGCTGGTGGTTTCAATAGTCTCCTCCTCAGCAGATC 1920
Db 1861 TTACACAAACAGCACCATGATGTGTGCTGGTGGTTTCAATAGTCTCCTCCTCAGCAGATC 1920
QY 1921 CTGCTATTCAGCTGGAAACAGTGTGATCGGCATCGGAGTGAAGCGCTGTTTACAGCA 1980
Db 1921 CTGCTATTCAGCTGGAAACAGTGTGATCGGCATCGGAGTGAAGCGCTGTTTACAGCA 1980
QY 1981 GGACCTGTATTCGCTGTGTGGAAACACAGGTGCTCAGTGTATCTCGTGGCGCTG 2040
Db 1981 GGACCTGTATTCGCTGTGTGGAAACACAGGTGCTCAGTGTATCTCGTGGCGCTG 2040
QY 2041 GCAACTGTATGAACAAAGAAAGATTAATAATCAGAAATGTTTTTCCAAAGAAATCTTTGAC 2100
Db 2041 GCAACTGTATGAACAAAGAAAGATTAATAATCAGAAATGTTTTTCCAAAGAAATCTTTGAC 2100
QY 2101 CATGACAGATGTGACCAACACAGATTTGTACAGCTGTACAGCCACACCAATGACTGC 2160
Db 2101 CATGACAGATGTGACCAACACAGATTTGTACAGCTGTACAGCCACACCAATGACTGC 2160
QY 2161 CACTGGTGAATGACCATTTGTGTCCTCCAGGAACACACAGCTGCTCAGAGGCCAGATCTCC 2220
Db 2161 CACTGGTGAATGACCATTTGTGTCCTCCAGGAACACACAGCTGCTCAGAGGCCAGATCTCC 2220
QY 2221 ATTTTGTAGTATGAGAAATGCCCCAGGATTAACCCCATGTACTGTATACAGAGACC 2280
Db 2221 ATTTTGTAGTATGAGAAATGCCCCAGGATTAACCCCATGTACTGTATACAGAGACC 2280
QY 2281 AGCTCAGAGCTGTGCTCCAGCAGAACTGCGAGTGGAGCCCGGAATCAGAGTGC 2340
Db 2281 AGCTCAGAGCTGTGCTCCAGCAGAACTGCGAGTGGAGCCCGGAATCAGAGTGC 2340
QY 2341 ATTTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCAATTTGGTGGAAATCTATGTTG 2400
Db 2341 ATTTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCAATTTGGTGGAAATCTATGTTG 2400
QY 2401 AAAATTAATCTACTGCAAGGAAATATGCAATGCTTAATTTGTTCTGTAGGAAACCAAT 2460
Db 2401 AAAATTAATCTACTGCAAGGAAATATGCAATGCTTAATTTGTTCTGTAGGAAACCAAT 2460
QY 2461 GCCCTTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTGCGA 2520
Db 2461 GCCCTTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTGCGA 2520
QY 2521 ATAATGCAATCTCTCAGAGCATGTCAGCTCACCITTAACCCCATGGGTGGCCCTTCGG 2580

Db 2521 ATAATGCAATCTCTCAGAGCATGTCAGCTCACCITTAACCCCATGGGTGGCCCTTCGG 2580
QY 2581 AAGATCAAATGTCTCTACTGCTGGGAGATATGTCCTCCCATTTACAAATAGTTTACTA 2640
Db 2581 AAGATCAAATGTCTCTACTGCTGGGAGATATGTCCTCCCATTTACAAATAGTTTACTA 2640
QY 2641 CAGTGAATGCCGTCTGAGCCAGTGTGATGCTGTGGAATTTTATCAGAACCCAGT 2700
Db 2641 CAGTGAATGCCGTCTGAGCCAGTGTGATGCTGTGGAATTTTATCAGAACCCAGT 2700
QY 2701 ACTCGGGGACTGAAGCTGCAACCTGCAACCACTCAATGCTGTGTGAAAGG 2760
Db 2701 ACTCGGGGACTGAAGCTGCAACCTGCAACCACTCAATGCTGTGTGAAAGG 2760
QY 2761 CTTGCAAAACACAGTCTAAGCAGTGCAGCACCATGTCCTTGAGGACAGCATGTGGA 2820
Db 2761 CTTGCAAAACACAGTCTAAGCAGTGCAGCACCATGTCCTTGAGGACAGCATGTGGA 2820
QY 2821 GATTGACACAGCGGAGCTCTGAGTGCATGCTGGTGCAGCAACATGAAGCAGTGTGGAC 2880
Db 2821 GATTGACACAGCGGAGCTCTGAGTGCATGCTGGTGCAGCAACATGAAGCAGTGTGGAC 2880
QY 2881 TCCAAATGCTATGTGGCCCTCCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAC 2940
Db 2881 TCCAAATGCTATGTGGCCCTCCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAC 2940
QY 2941 ACCTGCCCCCTGAAAATTTGTTTCCAGGCTACTGTACCTGTAGTCAITTTGTTGAGCAACCA 3000
Db 2941 ACCTGCCCCCTGAAAATTTGTTTCCAGGCTACTGTACCTGTAGTCAITTTGTTGAGCAACCA 3000
QY 3001 GGCTGTGGTGTGTAATGCTATCCAGCAATCTGCAAGGAAATGCAATAGAGGTTCC 3060
Db 3001 GGCTGTGGTGTGTAATGCTATCCAGCAATCTGCAAGGAAATGCAATAGAGGTTCC 3060
QY 3061 TATAAAGGACCAAGTGAAGTGTCTCGAAGCCCTTACAGGAAATTTCTATCCACAGCC 3120
Db 3061 TATAAAGGACCAAGTGAAGTGTCTCGAAGCCCTTACAGGAAATTTCTATCCACAGCC 3120
QY 3121 CTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATCAATGCTGCTTCTTCACTCT 3180
Db 3121 CTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATCAATGCTGCTTCTTCACTCT 3180
QY 3181 CCAGCTTGCCAAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGT 3240
Db 3181 CCAGCTTGCCAAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGT 3240
QY 3241 GAGAACCTGACACAGGCAAGCACTGCGAGACCTGCAATCTGCTTCTACCGTGTATCC 3300
Db 3241 GAGAACCTGACACAGGCAAGCACTGCGAGACCTGCAATCTGCTTCTACCGTGTATCC 3300
QY 3301 ACCAATGGAGGAAATGTCAGCCATGCAATGCAATGGCAGCGCTCTGTGCAACAC 3360
Db 3301 ACCAATGGAGGAAATGTCAGCCATGCAATGCAATGGCAGCGCTCTGTGCAACAC 3360
QY 3361 AACACGGGCAAGTGTCTTGTGACCAACAGGGCGTCAAGGGGACAGTGCAGCTATGT 3420
Db 3361 AACACGGGCAAGTGTCTTGTGACCAACAGGGCGTCAAGGGGACAGTGCAGCTATGT 3420
QY 3421 GAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTTATCTCTTCT 3480
Db 3421 GAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTTATCTCTTCT 3480
QY 3481 ATTGACTATCAGTTCACCTTTTACTTATCCAGGAAAGATGATCGCTATTTACACAGTATC 3540
Db 3481 ATTGACTATCAGTTCACCTTTTACTTATCCAGGAAAGATGATCGCTATTTACACAGTATC 3540
QY 3541 AATTTTGTGGTACTCTCTGACGAAACAAACAGGGATTTGGACATGTTTCAATATGCTCC 3600
Db 3541 AATTTTGTGGTACTCTCTGACGAAACAAACAGGGATTTGGACATGTTTCAATATGCTCC 3600
QY 3601 AAGATTTTCAACTCAACATCACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660

Db 1560 GCCTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGATAC 1619
Qy 1623 CCAGATGTGGACCAATCTTAAAGGACAGCCGATTTTTCGGTTACTTGGACACAGCTGTGAT 1682
Db 1620 CCAGATGTGGACCAATCTTAAAGGACAGCCGATTTTTCGGTTACTTGGACACAGCTGTGAT 1679
Qy 1683 AGTGAATGGAAACCAATGCTGGTGTGTTGGGGGAAACACACACATGACATCTATGAGCCA 1742
Db 1680 AGTGAATGGAAACCAATGCTGGTGTGTTGGGGGAAACACACACATGACATCTATGAGCCA 1739
Qy 1743 TGGCCCAAAATGCTTCTCTTCAGATTTTCATGGCCCTATGACATTTGCTGTGACCGCTGTGC 1802
Db 1740 TGGCCCAAAATGCTTCTCTTCAGATTTTCATGGCCCTATGACATTTGCTGTGACCGCTGTGC 1799
Qy 1803 AGTGTCTCCAGACCTGATCTCCACCATGATGTGTAACAGATTTGGCCATTCAGACATCTT 1862
Db 1800 AGTGTCTCCAGACCTGATCTCCACCATGATGTGTAACAGATTTGGCCATTCAGACATCTT 1859
Qy 1863 ACACAAACAGCACCATGATGTGTTGCGTGGTTCCTCAATAGTCTCTCTCCAGCACATCCT 1922
Db 1860 ACACAAACAGCACCATGATGTGTTGCGTGGTTCCTCAATAGTCTCTCTCCAGCACATCCT 1919
Qy 1923 GGTATTCACCTCGGAACAGTGTGATGCGCATGGAGTGAAGCCGCTGTTTATGACAGAGG 1982
Db 1920 GGTATTCACCTCGGAACAGTGTGATGCGCATGGAGTGAAGCCGCTGTTTATGACAGAGG 1979
Qy 1983 ACCTGGTATTCGGTGTGTGGAAACACAGGTCCTCTCAGTGTATCTCGTGGGCGCTGGC 2042
Db 1980 ACCTGGTATTCGGTGTGTGGAAACACAGGTCCTCTCAGTGTATCTCGTGGGCGCTGGC 2039
Qy 2043 AACTGATGAACAAAGAAAGAAAGTAAAATCAGAATGTTTTTCCAAAAGAACTTTGACCA 2102
Db 2040 AACTGATGAACAAAGAAAGAAAGTAAAATCAGAATGTTTTTCCAAAAGAACTTTGACCA 2099
Qy 2103 TGACAGATGTGACCAAGCAGACAGATTTGTAAGCTGTGACAGCAACACCAATGATGCA 2162
Db 2100 TGACAGATGTGACCAAGCAGACAGATTTGTAAGCTGTGACAGCAACACCAATGATGCA 2159
Qy 2163 CTGTGTCAATGACATTTGTGTCCTCCAGAAACACAGCTGCTCAGAAAGCCAGATCTCCAT 2222
Db 2160 CTGTGTCAATGACATTTGTGTCCTCCAGAAACACAGCTGCTCAGAAAGCCAGATCTCCAT 2219
Qy 2223 TTTTAGGTATGAGATTTGCCCAAGATAAOCCTATGTAATGTAATCAAGAGAACAG 2282
Db 2220 TTTTAGGTATGAGATTTGCCCAAGATAAOCCTATGTAATGTAATCAAGAGAACAG 2279
Qy 2283 CTGAGGAGCTGTGCTGACAGCACTGCGAGTGGGAGCCCGGATCAGGATGCAAT 2342
Db 2280 CTGAGGAGCTGTGCTGACAGCACTGCGAGTGGGAGCCCGGATCAGGATGCAAT 2339
Qy 2343 TGCCCTGCCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGAAACTCATGTTTCAA 2402
Db 2340 TGCCCTGCCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGAAACTCATGTTTCAA 2399
Qy 2403 AATTACTACTGCCAAGGAGAAATATGACATGCTAAATTTGTTCTGTAGGAACCAATGC 2462
Db 2400 AATTACTACTGCCAAGGAGAAATATGACATGCTAAATTTGTTCTGTAGGAACCAATGC 2459
Qy 2463 CTTTTTGGCTTCTCTTACAAACCCAGAAAGTGTAGAAATTTGTCCTTAAGCAGCTCGGAAT 2522
Db 2460 CTTTTTGGCTTCTCTTACAAACCCAGAAAGTGTAGAAATTTGTCCTTAAGCAGCTCGGAAT 2519
Qy 2523 AATGAGTCAATCTCAGAGCATGTCCAAAGCTCACTTAAACCCCATCGGTCGGCTTCGAA 2582
Db 2520 AATGAGTCAATCTCAGAGCATGTCCAAAGCTCACTTAAACCCCATCGGTCGGCTTCGAA 2579
Qy 2583 GATCAATGTCTACTGCTGGGAGATATGCCCCATTTACAAATAGTTTACTACA 2642
Db 2580 GATCAATGTCTACTGCTGGGAGATATGCCCCATTTACAAATAGTTTACTACA 2639
Qy 2643 GTGGATGCTGCTGAGCCCGAGTGTGCTGGAATTTCTGTGGAATTTTATCAGAACCCAGTAC 2702
Db 2640 GTGGATGCTGCTGAGCCCGAGTGTGCTGGAATTTCTGTGGAATTTTATCAGAACCCAGTAC 2699

Qy 2703 TCGGGACCTGAAGGCTGCAACCTGTGATCAACCCACTCAATGGTAGTGTCTGTGAAGGCC 2762
Db 2700 TCGGGACCTGAAGGCTGCAACCTGTGATCAACCCACTCAATGGTAGTGTCTGTGAAGGCC 2759
Qy 2763 TGCAAACCAACAGTGTAAAGCAGTGCAGGACACCATGTGCTTTCAGGACAGCATGTGAGA 2822
Db 2760 TGCAAACCAACAGTGTAAAGCAGTGCAGGACACCATGTGCTTTCAGGACAGCATGTGAGA 2819
Qy 2823 TTAGCAGGCGGAGCTCTGAGTGTGATGTGCTGAGCAACCATGTGCTTTCAGGACAGCATGTGAGTCC 2882
Db 2820 TTAGCAGGCGGAGCTCTGAGTGTGATGTGCTGAGCAACCATGTGCTTTCAGGACAGCATGTGAGTCC 2879
Qy 2883 CAATGCTATGTGCGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATATACATGAGCAC 2942
Db 2880 CAATGCTATGTGCGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATATACATGAGCAC 2939
Qy 2943 CTGCCCCCTGAAAAATGTGTTTTCAGGCTACTGTACCTGTAGTCAATTTGCTGGAGCAACAGG 3002
Db 2940 CTGCCCCCTGAAAAATGTGTTTTCAGGCTACTGTACCTGTAGTCAATTTGCTGGAGCAACAGG 2999
Qy 3003 CTGTGGCTGTGTACTGATCCCGAGCAATACTGCAAGGGAATGATGAGGGTTCCTA 3062
Db 3000 CTGTGGCTGTGTACTGATCCCGAGCAATACTGCAAGGGAATGATGAGGGTTCCTA 3059
Qy 3063 TAAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122
Db 3060 TAAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119
Qy 3123 GCTCAATTCAGCATGTGTCTGAGGACAGCAGATCAACTGTGCTTTTCAATCAGTGTCC 3182
Db 3120 GCTCAATTCAGCATGTGTCTGAGGACAGCAGATCAACTGTGCTTTTCAATCAGTGTCC 3179
Qy 3183 AGCTTGCAATGCAAGCCGACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGA 3242
Db 3180 AGCTTGCAATGCAAGCCGACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGA 3239
Qy 3243 GAACCTGACCAAGGCAAGCACTGCGAGACCTGCATATCTGCTTTCAGCGTATCCAC 3302
Db 3240 GAACCTGACCAAGGCAAGCACTGCGAGACCTGCATATCTGCTTTCAGCGTATCCAC 3299
Qy 3303 CAATGAGGAAATGTGAGCCATGCAAGTGCATGCGGACCGCTCTCTGTGCAACCA 3362
Db 3300 CAATGAGGAAATGTGAGCCATGCAAGTGCATGCGGACCGCTCTCTGTGCAACCA 3359
Qy 3363 CAGGGCAAGTGTCTTTCGACACCAAGGCGTCAAGGGGAGCAGTGCAGCTATGTA 3422
Db 3360 CAGGGCAAGTGTCTTTCGACACCAAGGCGTCAAGGGGAGCAGTGCAGCTATGTA 3419
Qy 3423 GGTAGAAATCGATAACAGGAAACCTCTCAGAGGAAACATGTTATTTATCTTCTTAT 3482
Db 3420 GGTAGAAATCGATAACAGGAAACCTCTCAGAGGAAACATGTTATTTATCTTCTTAT 3479
Qy 3483 TGACTATCAGTTTCACTTTTGTATCTCCAGGAAGATGATCGCTATTTACAGCTATCAA 3542
Db 3480 TGACTATCAGTTTCACTTTTGTATCTCCAGGAAGATGATCGCTATTTACAGCTATCAA 3539
Qy 3543 TTTTGTGGCTACTCTGACGAAACAAACAGGATTTGGACATGTTTCAATCAATGCTCCAA 3602
Db 3540 TTTTGTGGCTACTCTGACGAAACAAACAGGATTTGGACATGTTTCAATCAATGCTCCAA 3599
Qy 3603 GAATTTCAAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGAGA 3662
Db 3600 GAATTTCAAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGAGA 3659
Qy 3663 AGAGATGCTGTTTGTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA 3722
Db 3660 AGAGATGCTGTTTGTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA 3719
Qy 3723 GAAGTTTGAATTTTCGCAACCCCAATATCACTTTCTTTTGTGTTATGTCAGTAAATTCAC 3782
Db 3720 GAAGTTTGAATTTTCGCAACCCCAATATCACTTTCTTTTGTGTTATGTCAGTAAATTCAC 3779

QY 1743 TGGGCCAAATGCTTCTCTCAGATTTTCATGGCCCTATGACATTCCTGTGACCGCTGGTC 1802
Db 1740 TGGGCCAAATGCTTCTCTCAGATTTTCATGGCCCTATGACATTCCTGTGACCGCTGGTC 1799
QY 1803 AGTGGTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGAGTCTT 1862
Db 1800 AGTGGTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGAGTCTT 1859
QY 1863 ACACAACAGCACCATGATGTGTTGGTGGTGGTTCCTCAATAGTCTCCTCCTCAGCAGATCCT 1922
Db 1860 ACACAACAGCACCATGATGTGTTGGTGGTGGTTCCTCAATAGTCTCCTCCTCAGCAGATCCT 1919
QY 1923 GGTATTCACCTCGGACAGTGTGATGCGCATCGGAGTGAAGCCGCTGGTTCAGCAGAG 1982
Db 1920 GGTATTCACCTCGGACAGTGTGATGCGCATCGGAGTGAAGCCGCTGGTTCAGCAGAG 1979
QY 1983 ACCTGGTATTCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGGCGCTGGC 2042
Db 1980 ACCTGGTATTCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGGCGCTGGC 2039
QY 2043 AACTGATGAACAAAGAAAGTAAATCAAGATGTTTTTCCAAAAGAACTCTTGACCA 2102
Db 2040 AACTGATGAACAAAGAAAGTAAATCAAGATGTTTTTCCAAAAGAACTCTTGACCA 2099
QY 2103 TGACAGATGTGACAGACACAGATTTGTTACAGCTGTACAGCCACCAATGACTGCCA 2162
Db 2100 TGACAGATGTGACAGACACAGATTTGTTACAGCTGTGACAGCCACCAATGACTGCCA 2159
QY 2163 CTGTGCAATGACATTTGTCCCAGGAACACAGCTGCTCAGAAGGCCAGATCTCCAT 2222
Db 2160 CTGTGCAATGACATTTGTCCCAGGAACACAGCTGCTCAGAAGGCCAGATCTCCAT 2219
QY 2223 TTTTAGTATGAGATTTCCCGCAGGATAACCCCATGTACTACTGTAAACAGAGACAG 2282
Db 2220 TTTTAGTATGAGATTTCCCGCAGGATAACCCCATGTACTACTGTAAACAGAGACAG 2279
QY 2283 CTGACAGAGCTGTGCCCTGGACCAAGAACTGCCAGTGGGAGCCCGGAACTCAGGAGTGCAT 2342
Db 2280 CTGACAGAGCTGTGCCCTGGACCAAGAACTGCCAGTGGGAGCCCGGAACTCAGGAGTGCAT 2339
QY 2343 TGCCCTGCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCATGTTGAA 2402
Db 2340 TGCCCTGCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCATGTTGAA 2399
QY 2403 AATTACTACTGCCAGGAGAAATTATGACATGCTAAATTTGTTCTGTAGAACCAACATGC 2462
Db 2400 AATTACTACTGCCAGGAGAAATTATGACATGCTAAATTTGTTCTGTAGAACCAACATGC 2459
QY 2463 CTTTTTGGCTTCTCTTACAAACCAAGAGAGTAGAATTTGTCTTAAAGCAGCTGCGAAT 2522
Db 2460 CTTTTTGGCTTCTCTTACAAACCAAGAGAGTAGAATTTGTCTTAAAGCAGCTGCGAAT 2519
QY 2523 AATGAGTCACTCAGAGATGTCCAGCTCAGCTTAACCCCATGGTGGGCTTCGGAA 2582
Db 2520 AATGAGTCACTCAGAGATGTCCAGCTCAGCTTAACCCCATGGTGGGCTTCGGAA 2579
QY 2583 GATCAATGTCTCTACTGTGTGGAGAGATATGCCCAATTTACAAATAGTTACTACA 2642
Db 2580 GATCAATGTCTCTACTGTGTGGAGAGATATGCCCAATTTACAAATAGTTACTACA 2639
QY 2643 GTGGATGCGGTGTGAGCCAGTGTGATGCTGGATTTGTGGAAATTTATCAGAACCCAGTAC 2702
Db 2640 GTGGATGCGGTGTGAGCCAGTGTGATGCTGGATTTGTGGAAATTTATCAGAACCCAGTAC 2699
QY 2703 TCGGGGACTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTCTGTGAAAGGCC 2762
Db 2700 TCGGGGACTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTCTGTGAAAGGCC 2759
QY 2763 TGCAAAACCAAGTGTAGAGCTGCGGACACCAATGTGCTTTGAGGACAGCATGTGGAGA 2822
Db 2760 TGCAAAACCAAGTGTAGAGCTGCGGACACCAATGTGCTTTGAGGACAGCATGTGGAGA 2819
QY 2823 TTGACACAGCGCAGCTCTGAGTGTGATGTGGTGACGACATGAAGCAGTGTGGACTC 2882

RESULT 4
AF034957

LOCUS

DEFINITION Homo sapiens secreted T cell activation protein Attractin
(attractin) mRNA, complete cds.

3597 bp mRNA linear

PRI 30-SEP-1998

Db 2820 TTGCACACGCGCAGCTCTGAGTGCATGTGGTGACCAACATGAACAGTGTGTGACTC 2879
QY 2883 CAATGCTATGTGGCCCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGCAC 2942
Db 2880 CAATGCTATGTGGCCCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGCAC 2939
QY 2943 CTGCCCCCTGAAAAATTTGTTTCAAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACAGG 3002
Db 2940 CTGCCCCCTGAAAAATTTGTTTCAAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACAGG 2999
QY 3003 CTGTGGTGTGTGATGATCCAGCAATATCTGGCAAGGGAATCATAGAGGGTTCCTA 3062
Db 3000 CTGTGGTGTGTGATGATGATCCAGCAATATCTGGCAAGGGAATCATAGAGGGTTCCTA 3059
QY 3063 TAAAGACACAGTGAAGATGCTTTCGCAAGCCCTACAGAAATTTTATATCCAGAGCCCT 3122
Db 3060 TAAAGACACAGTGAAGATGCTTTCGCAAGCCCTACAGAAATTTTATATCCAGAGCCCT 3119
QY 3123 GCTCAATTCAGCATGTGTCTAGAGACAGCAGATACATCTGCTTTTTCATTCACATGCTC 3182
Db 3120 GCTCAATTCAGCATGTGTCTAGAGACAGCAGATACATCTGCTTTTTCATTCACATGCTC 3179
QY 3183 AGCTTGCATATGCAACGCGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGA 3242
Db 3180 AGCTTGCATATGCAACGCGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGA 3239
QY 3243 GAACCTGACACAGGACAGCATCTGCAGACCTGTCATATCTGCTTTTCTACGCTGATCCCAC 3302
Db 3240 GAACCTGACACAGGACAGCATCTGCAGACCTGTCATATCTGCTTTTCTACGCTGATCCCAC 3299
QY 3303 CAATGAGGGAATGTGAGCCATGCAAGTGCATGCGGACCGCTCTCTGTGCAACACCAA 3362
Db 3300 CAATGAGGGAATGTGAGCCATGCAAGTGCATGCGGACCGCTCTCTGTGCAACACCAA 3359
QY 3363 CACGGCAAGTGTCTGTGACCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGA 3422
Db 3360 CACGGCAAGTGTCTGTGACCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGA 3419
QY 3423 GGTAGAAATTCGATACCAAGAAACCTCTCAGAGGAACATGTTATATATCTCTTCTTAT 3482
Db 3420 GGTAGAAATTCGATACCAAGAAACCTCTCAGAGGAACATGTTATATATCTCTTCTTAT 3479
QY 3483 TGAATATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATCAA 3542
Db 3480 TGAATATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATCAA 3539
QY 3543 TTTTGTGGCTACTCTGACGAAACAAACAGGGATTTGGACATGTTTATCAATGCTCCAA 3602
Db 3540 TTTTGTGGCTACTCTGACGAAACAAACAGGGATTTGGACATGTTTATCAATGCTCCAA 3599
QY 3603 GAATTTCAACCTCAACATCACCTGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGGAGA 3662
Db 3600 GAATTTCAACCTCAACATCACCTGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGGAGA 3659
QY 3663 AGAGATGCTGTGTTTTCAAAAACCAACATTAAGGAGTCAAGAGATGTTTCTCTAATGA 3722
Db 3660 AGAGATGCTGTGTTTTCAAAAACCAACATTAAGGAGTCAAGAGATGTTTCTCTAATGA 3719
QY 3723 GAAGTTTGAATTTTGGCAACCCACCAATATCATTCTTTTGTATGTCAGTAAATTCAC 3782
Db 3720 GAAGTTTGAATTTTGGCAACCCACCAATATCATTCTTTTGTATGTCAGTAAATTCAC 3779
QY 3783 CTGGCCCATCAAAATTCAGGT 3803
Db 3780 CTGGCCCATCAAAATTCAGAT 3800

ACCESSION AF034957
 VERSION AF034957.1 GI:3676346
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3597)
 AUTHORS Duke-Cohan, J.S., Gu, J., McLaughlin, D.F., Xu, Y., Freeman, G.J. and Schlozman, S.F.
 TITLE Attractin (DPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
 MEDLINE 98409658
 PUBMED 9736737
 REFERENCE 2 (bases 1 to 3597)
 AUTHORS Duke-Cohan, J.S., Gu, J., Ao, Z., McLaughlin, D.F., Freeman, G.J. and Schlozman, S.F.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
 FEATURES
 source
 gene
 CDS
 Location/Qualifiers
 1..3597
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..3597
 /gene="attractin"
 /note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; DPT-L"
 /codon_start=1
 /product="Attractin"
 /protein_id="AAC61902.1"
 /db_xref="GI:3676347"
 /translation="MVAAATEARLRETAALAGSGGPHVNGRCNPCTGOC VCPAGVGEQCHCGFRLTGSSGFVDGNGYKTKCTWLEGGPNRLRLRPH FATECSWHLVYDGSIIYAPLVAAPSLVPERDNETPEVVAISYALHPFSDA AYNGFNITVDFMCPNKGSGECKISNSDVECESENWKEACDIPHCYDNG PFHIGICNSDVGGSCFSQWQPGCVPEPANQSFRTREYSNLKLPRAASHAVNG NIMVYGGVMFNHSDYMLAYDLASREWLPLNRSVNNVVYVYGHSLYKDYIMYIG KIDSTGNVTNELVFPHTHNESWLLTPKAQVAVVGHSAHIVTLKNGRVMLVIFG HCLVYISNVQVEYDLKNTWLLHTQCALVQGGYGHSSVVDHRTALYVHGYSKFS ANKYRLADLLRYVDVTQWVLLKDSRFRYLHTAVVSGTMLVFGNTHNDTMSHG AKPSDFMAYDIADRSVLPDPSTWMSLDLPAVLHNSMTWVFGFNLSLSDI LVFTSEQDHRSEAACLAAGPGRVWNTSSQCI SWALATDEQBEKLKSECKRT LDHRCQDHTCYSTANTNDCHNDHCVPRNSCSGQISIPYENCNDNPMYIC NKTKSCRSALDQNCQWEPNQEGLALPENICGHWLVGNLSKLTAKENYDNAL FCRHNALLASLTOKKVEFVLKOLRIMOSSQMSKLTLPWGLRKNVSYWCWDM SPETNSLQWPSRPSDAGFCGILSEPSTRGLKATCINPLNGSVCRERPAHSAKQCF TPCALRTACGCTGSSSCMCSNWKQCVDSNAYVASFPFGQCMEWTMSCTCPENCS GYCTCSHCEQPCWCNDPSTNGKRCIEGYSYKGPVMPSPQAPTGNFYQPLNLSM CLEDSRYNHSFHCACQCNHSCIKINQSI CEKCNLTGKHCETCISGFYDPTNGG KCQPCKNYGHASLNTNTGKCFCTTKGVDCEQLCEVKNRYQGNPLRGCTCYTLLID YQFTSLSQEDDRYTAINFVATPDQNRDLDFINASKNFNLNITWAAFSFSAQTQAG BEMPVVSXTNIKKEYKDSFNEKDFRNHPNITFFVYVSNTFWPIKIQVTEQ"

ORIGIN

Query Match 91.3%; Score 3487; DB 9; Length 3597;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;

QY 270 GTCGGGCTCAGCGCAGCGGAGCCAGCAAGAAATGTGACCGGCCCTGTGTCAACGCGGTCG 329
 DB 48 GCGCGGCGAGCGAGCGCTCGCGGGCAGGAGCGGCGCGGCGGCTGTGTCAACGCGGTCG 107
 QY 330 CTGCAACCTGCGCACCGGCGAGTGTGCTGTGCCCCCGCGGCTGGGTGGCGGAGCAATGCCA 389
 DB 108 CTGCAACCTGCGCACCGGCGAGTGTGCTGTGCCCCCGCGGCTGGGTGGCGGAGCAATGCCA 167

QY 390 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGTTTGTGACAGATGCACTGG 449
 DB 168 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGTTTGTGACAGATGCACTGG 227
 QY 450 AAATTATAATACAAACGAAGTGCACGCGCTCATTTGAAGGACAGCCAAATAGATAAT 509
 DB 228 AAATTATAATACAAACGAAGTGCACGCGCTCATTTGAAGGACAGCCAAATAGATAAT 287
 QY 510 GAGACTTCGTTTCAATCAATTTCTCTACAGAGTGTAGTTGGACCAATTTATATGTTATGA 569
 DB 288 GAGACTTCGTTTCAATCAATTTCTCTACAGAGTGTAGTTGGACCAATTTATATGTTATGA 347
 QY 570 TGGGACTCAATTTATGACCCGCTAGTCTGCTGCAATTTAGTGGCTCATTTGTTCCCTGAGAG 629
 DB 348 TGGGACTCAATTTATGACCCGCTAGTCTGCTGCAATTTAGTGGCTCATTTGTTCCCTGAGAG 407
 QY 630 AGATGGCAATGAGACTGTCCCTGAGCTTGTGGACATCAGGTTATGCCCTTGTGCTGCAATTT 689
 DB 408 AGATGGCAATGAGACTGTCCCTGAGCTTGTGGACATCAGGTTATGCCCTTGTGCTGCAATTT 467
 QY 690 TTTTAGTGATGCTGCTTATAATTTGACTGATTTAATATTACTTACAGTTTGTATGTTG 749
 DB 468 TTTTAGTGATGCTGCTTATAATTTGACTGATTTAATATTACTTACAGTTTGTATGTTG 527
 QY 750 TCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGTCAAGTAAATAGCAGGAACTGTTGA 809
 DB 528 TCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGTCAAGTAAATAGCAGGAGTACTGTTGA 587
 QY 810 ATGTGAATGTTTGAATACTTGAAGGTGAAGCATGTGATTCCTCACTGTACAGACAA 869
 DB 588 ATGTGAATGTTTGAATACTTGAAGGTGAAGCATGTGATTCCTCACTGTACAGACAA 647
 QY 870 CTGTGTTTTCCTCATCGAGGATCTGCAATTCAGTGTAGTGTACAGGATGCTCCTGCTT 929
 DB 648 CTGTGTTTTCCTCATCGAGGATCTGCAATTCAGTGTAGTGTACAGGATGCTCCTGCTT 707
 QY 930 CTCAGACTGCGGCTCCTGGATGTTTCAGTTCCTGTACAGCTAACCACTGATTTTGGAC 989
 DB 708 CTCAGACTGCGGCTCCTGGATGTTTCAGTTCCTGTACAGCTAACCACTGATTTTGGAC 767
 QY 990 TCAGAGGAATATTTAACTTAAAGTCCCAGAGCATCTATAAGCTGTGGTCAATGG 1049
 DB 768 TCAGAGGAATATTTAACTTAAAGTCCCAGAGCATCTATAAGCTGTGGTCAATGG 827
 QY 1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGTTCT 1109
 DB 828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGTTCT 887
 QY 1110 AGCGTATGACTTGTCTTAGGAGTGGCTTCCATTAACCGTCTGTGTGAACAATGTGT 1169
 DB 888 AGCGTATGACTTGTCTTAGGAGTGGCTTCCATTAACCGTCTGTGTGAACAATGTGT 947
 QY 1170 TGTTAGATATGTCATTTCTTTTGCATTATACAAGGATAAAATTTACATGTATGGAGAAA 1229
 DB 948 TGTTAGATATGTCATTTCTTTTGCATTATACAAGGATAAAATTTACATGTATGGAGAAA 1007
 QY 1230 AATTGATCAACTGGGAATGTGACCAATGAGTGTGAGTGTTCATCAATTCATTAATGAGTC 1289
 DB 1008 AATTGATCAACTGGGAATGTGACCAATGAGTGTGAGTGTTCATCAATTCATTAATGAGTC 1067
 QY 1290 ATGGGTGTTGTTGACCCCTTAAGGCAAGGACAGTGTGAGTGTGGGCACTCTGCACA 1349
 DB 1068 ATGGGTGTTGTTGACCCCTTAAGGCAAGGACAGTGTGAGTGTGGGCACTCTGCACA 1127
 QY 1350 CATTTGTACACTGAAGATGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409
 DB 1128 CATTTGTACACTGAAGATGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187
 QY 1410 CTATGATATATAGCAATGTGAGGAATATGATTTGATTAAGAACACATGAGTATATT 1469
 DB 1188 CTATGATATATAGCAATGTGAGGAATATGATTTGATTAAGAACACATGAGTATATT 1247


```

Db      3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3466
QY      3689 ACATTAAGAGATCAAAAGATGTTTCTTAATGAGAAGTTGATTTTCGCAACCAACCAA 3748
Db      3467 ACATTAAGAGATCAAAAGATGTTTCTTAATGAGAAGTTGATTTTCGCAACCAACCAA 3526
QY      3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTTACCTTGCCCAATCAAAATTCAGGTGCAA 3808
Db      3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTTACCTTGCCCAATCAAAATTCAGGTGCAA 3586
QY      3809 CTGAACAATGA 3819
Db      3587 CTGAACAATGA 3597

RESULT 5
AF531101
LOCUS      4493 bp mRNA linear MAM 19-AUG-2002
DEFINITION Bos taurus attractin mRNA, complete cds.
ACCESSION AF531101
VERSION AF531101.1 GI:22297307
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 4493)
Graphodatskaya, D., Joerg, H. and Stranzinger, G.
Direct Submission
Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,
Zurich, CH 8092, Switzerland
Location/Qualifiers
1. 4493
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/chromosome="13"
52. 4299
/note="transmembrane form; mahogany protein"
/codon_start=1
/product="attractin"
/protein_id="AA095445.1"
/db_xref="GI:22297308"
/translation="MVAATEARLRIRVTVTPAPTSRNGRRCDWDATRAGRPLRGL
RLPRRRLILLPLLLLPAAEAAAASVSGSAAYKTKCTWLIEGPNRIMRLR
GQCVPAGWGEQCHQGRFLTGSGFITDGPNGYKTKCTWLIEGPNRIMRLR
FNHPATECSWHLVYDGDLSIYAPLVAFAFGLIYPERDGNETPEVATSGYALLHFF
SDAAYNLGNITTYNFMCPNCSGRGCEKLSNSNTVQCESNWKGEACDIPHCYN
NGPFHRCIGNSDVRGCSFSEWQPGCVFPANQSFWTREYSDLKLPRAHAKV
VNGIMVVGVSNEHNSDYNMAYDLASREWLALNSVNSVYVRYGHSALYODKIY
MYGKIDSTGNVRELRYHINHSWLLSPKAEQYAVVGHSAHIVTLKNGRVMLV
IFGHCPLYGYSVQVEYDLKNTWILITGALVQGGYGHSSVYDDTKALYIHGGYK
AFSANKYRLADLVYDVTQMTWILKDSRFRYLHTAVIVSGTMLVFGGNTNDTSM
SHGAKRDFMAVIDDCAWVLPRLDHDVNRFGHSAVLNRYMVFSGNSLL
SDILVFTSECEAQSENAACLAAGPVRCDMTDSSQCVSELAPEAQEKSECRSK
RIFDHRCDQITDICTANTANGCNCNDHCVFNHSCTEGQISFIYDHPKDNPMY
YCNKTSRCSALDONQWEPNRCIALPENICIGIWHLVGNSCLITARTYDNA
KLSCRNHNAFLASTITQKKEVFLKQRLIMQSSQMSKFTLTPWGLRLKLVNYSWQE
DMSPFTSLQWMESEPSDAGFCGILSEPSRGLKAATCINPLNGSVCEPANHSAQ
CRTPCALRTACGETSGSCECMWCKQCVDSNAYVAFSPFGOCMEWYMSCPDEN
SCGYCTSHCLEQPCGWCTDPSNTGKGIKESYKGVKPMQSGPTGNSYPOPLNS
SMKLEDSPYNSPFIHCPACQNGHSHKINSIGICEKNLITGKHCEFCISGFQDPN
GKQCPKRCNGHSLCNTNGKCTCTTGNGVDEQCEVENRYQGNPLAGTCTYILL
IDYQFTSLSQEDDRYYTAINFATPDQNRDLDMFINASKNLFNLTWAFSFSQNF
AGEEMPVVSKTNIKEYKDSFNSNEFDRPHNFNITFFVYNSNFTWPKIQIAPSHQNF
MDLVQFFVTFPSFCLSLVAAVVMKIQSCWASRRREQLLRMQQWASRPFASVNA
LETDEBPDLLTGSGIKVTPKPIALEPCFGNKAALVSFVRLPRGLGGIPPGQSGLA
ASALVDISQMPVVVYKESGAVNRKQPPAQPTCI"

```

ORIGIN

Query Match 85.1%; Score 3251.8; DB 4; Length 4493;
 Best Local Similarity 92.0%; Pred. No. 0;

```

Matches 3488; Conservative 0; Mismatches 267; Indels 36; Gaps 4;
QY      13 GCAGCGGCAACAGGCTGAGAGAGAGACCGCGCGAGCGGACAGCGTTCGCGGGC 72
Db      58 GCAGCGGCGACTGAGGCGAGGCT---GAGGAGACCGTGTGTGACACCGCGCCACGAGC 114
QY      73 AGGAGCGGCGGCGCGCACTGGGACTGGGAGCTGACACGAGGCTGGGAGCGCGGCTGGGG 132
Db      115 AGGAACGCGCAGGCGGTGTGGGACTGGGACCTAGAGCTGGAAGCGCGGCGCTGCGG 174
QY      133 GCGGCGCTGCGCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGGCTGCTGCTGCTG 192
Db      175 ACCGGGCTGCGCTCCACGGCGG-----TGGCGGCGCTGCTG 213
QY      193 CTGTTGTTGCTCCCGCGGCGGCTGTTGCTGCTGCTGCTGCTGCTGAGGCGGCGCGG 252
Db      214 CTGCTGCTGCTCCCGGC-----GCTGCTGCTACTTCCCTGGGCGGCTGAGGCGCGG 264
QY      253 GCGGCGGCGGCGGCGGTGTGCGGCTCAGCGGCGGAGGCGCAAGAAATGTGACCGGCC 312
Db      265 GCGGCGGCGGCTGCTGCTGCGGCTCAGCGGCGGAGGCGGCGCAAGAAATGTGACCGGCC 324
QY      313 TGTGTCAAGCGGCGTCTGCTCAACCTGGCAACCGGCGGCGGCTGCTGCTGCTGCTGCTG 372
Db      325 TGTGTCAAGCGGCGGCGGCTGCAACCTGGCAACCGGCGGCGGCTGCTGCTGCTGCTGCTG 384
QY      373 GTGGCGGAGCAATGCCAGCACTGCGGGGCGGCGCTTACAGACTAACTGGATCTTCTGGTTT 432
Db      385 GTGGCGGAGCAATGCCAGCACTGCGGGGCGGCGCTTACAGACTAACTGGATCTTCTGGTTT 444
QY      433 GTGACAGATGGACCTCGAAATATATAATAACAAACGAGTGCACGTGGCTGCTATTAAGA 492
Db      445 ATAACAGATGGACCTCGAAATATATAATAACAAACGAGTGCACGTGGCTGCTATTAAGA 504
QY      493 CAGCCAAATAGATAATAGACTTGGTTTCAATCACTTTTGTCTACAGAGTGTAGTTGGGAC 552
Db      505 CAGCCAAATAGATAATAGACTTGGTTTCAATCACTTTTGTCTACAGAGTGTAGTTGGGAC 564
QY      553 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTCATTTAGTTGGC 612
Db      565 CATTTATATGTTTATGACGGGACTCAATTTATGACCGCTAGTTGCTCATTTAGTTGGC 624
QY      613 CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCTCGAGGTTGTTGCGCCATCAGGT 672
Db      625 CTCATTGTTCTGAGAGAGACGCGCAATGAGACTGTCCTCGAGAGGTTGTTGCGCCATCAGGT 684
QY      673 TATGCTTCTGCTCATTTTTTTAGTGTGCTGTATTAATTTTACACTGGAATTAATTAAT 732
Db      685 TATGCTTCTGCTCATTTTTTTAGTGTGCTGTATTAATTTTACACTGGAATTAATTAAT 744
QY      733 TACAGTTTGTATGTGTCCAAATAAATGCTCAGCGCGGAGAGCTGTAAGATCAGTAAT 792
Db      745 TACAATTTTGACATGTGTCCAAATAAATGCTCAGCGCGGAGAGATTAAGATCAGTAAT 804
QY      793 ACAGCGAACTGTTCAATGTAATGTTTGTGAAAACCTGGAAGGTAAGCATGTGACATT 852
Db      805 AGAGCAACACTGTTCAATGTAATGTTTGTGAAAACCTGGAAGGTAAGCATGTGACATT 864
QY      853 CCTCACTGTACAGACAACTGTGTTTCTCATTCAGGCGCATCTGCAATTCAGTGTATGTC 912
Db      865 CCTCACTGTGTAATAAATGTTGTTTCTCATTCAGGCGCATCTGCAATTCAGTGTATGTC 924
QY      913 AGAGATGCTCTGCTTCTACACTGCGGAGGTCCTGTGATGTTTCAAGTTCTGTACAGCT 972
Db      925 AGAGATGCTCTGCTTCTACAGTGGCAAGGTCCTGTGATGTTTCTGTGTACAGCT 984
QY      973 AACCACTCAATTTTGCATCTCGAGAGAAATTTCTAACTTTAAAGCTCCCCAGAGCATCTCAT 1032
Db      985 AACCACTCAATTTTGCATCTCGAGAGAAATTTCTGACCTTAAGCTCCCCAGAGCATCTCAC 1044
QY      1033 AAGCTGTGGTCAATGGAAACATTAATGTTGGGTTGTTGGAGGATATATTTCAACACTCA 1092
Db      1045 AAGCTGTGGTCAATGGAAATATAATGTTGGGTTGTTGGCGGATATATTTCAACACTCA 1104

```

1093 GATTATACATGGTTCTAGCCTATGACCTTGCTTCTAGGAGTGGCTTCACTAAACCGT 1152
1105 GATTACAACTGGTTCTAGCCTATGACCTTGCTTCTAGGAGTGGCTTCACTAAACCGT 1164
1153 TCTGTGAACAATGGTTGTAGATATGCTGATCTTTGGCATATATCAAGGATAAAAT 1212
1165 TCTGTGAACAATGGTTGTAGATATGCTGATCTTTGGCATATATCAAGGATAAAAT 1224
1213 TACATGTATGAGGAAAAATGATCCAACTGGGAATGACCAATGATGAGTGGAGTTT 1272
1225 TACATGTATGAGGAAAAATGATCCAACTGGGAATGACCAATGATGAGTGGAGTTT 1284
1273 CACATTCATATGATGATGGTGTGACCTCTAGGCAAGGAGGAGGATGACAGTG 1332
1285 CATATTCATATGATGATGGTGTGACCTCTAGGCAAGGAGGAGGATGACAGTG 1344
1333 GTTGGGCACTCTGCACACATTTGACACTGAAGATGGCCGAGTGGTCACTGCTGATC 1392
1345 GTTGGGCACTCGGCGCACATTTGACACTGAAGATGGCCGAGTGGTCACTGCTGATC 1404
1393 TTTGGTCACTGCTGCTCTATGGATATATAGCAATGTGAGGATATGATTTGGATAAG 1452
1405 TTTGGTCACTGCTGCTCTATGGATATATAGCAATGTGAGGATATGATTTGGATAAG 1464
1453 AACACATGGAGTATATACACACCGAGTGGCTTGTGCAAGGGGTTACGGCCATAGC 1512
1465 AACACATGGAGTATATACACACTCAGGGTGGCTTGTGCAAGGGGTTACGGCCATAGC 1524
1513 AGTGTATACGACATAGGACACCGGCTCTATAGTTCATGTTGGTCAAAAGCTTTGAGT 1572
1525 AGTGTATAGATAGGACACCGGCTCTATAGTTCATGTTGGTCAAAAGCTTTGAGT 1584
1573 GCGAATAGTACGGCTGAGATGATCTTACCGATATGATGATGATGATGATGATGATG 1632
1585 GCGAATAAATACCGGCTGAGATGATCTTACAGATGATGATGATGATGATGATGATG 1644
1633 ACCATTCATTAAGGACAGCGATTTTTCGGTTACTTSCACACAGTGTGATGATGATG 1692
1645 ACCATTCATTAAGGACAGCGATTTTTCGGTTACTTSCACACAGTGTGATGATGATG 1704
1693 ACCATGCTGGTGTGGGGAACACACACATGACATCTATGAGCATGGGCGCAAA 1752
1705 ACCATGCTGGTGTGGGGAACACACACATGACATCTATGAGCATGGGCGCAAA 1764
1753 TGTCTCTTTCAGATTTATGCTGCTATGATGATGCTGCTGCTGCTGCTGCTGCTGCT 1812
1765 TGTCTCTTTCAGATTTATGCTGCTATGATGATGCTGCTGCTGCTGCTGCTGCTGCT 1824
1813 AGACCTGATCTCCACCATGATGACACAGATTTGGCCATTCAGAGTCTTACACACAGC 1872
1825 AGACCTGATCTCCACCATGATGACACAGATTTGGCCATTCAGAGTCTTACACACAGC 1884
1873 ACCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
1885 ACCATGATGATTTTGGTGGTGTATAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
1933 TCGGAACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1992
1945 TCGGAACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2004
1993 CGGTGTGTGGAACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2052
2005 CGGTGTGTGGAACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064
2053 CAAGAAGAAAGTTAAATCAGATGTTTTCAGAAAGAACTCTTGAACCATGACAGATGT 2112
2065 C---AAGAAAGAAATAGTCAAGATGTTTCTAAGAAATTTTGAACCATGACAGATGT 2121
2113 GACCAACACAGATTTGATGATGCTGATGACGCAACACCAATGATGCTGCTGCTGCTGCT 2172
2122 GACCACTACGATTTGATGATGCTGATGACGCAACACCAATGATGCTGCTGCTGCTGCTG 2181

2173 GACCATTTGTGCTCCAGGAACACACAGTCTGCTCAGAAAGGCGAGATCTCCATTTTAGGTAT 2232
2182 GACCATTTGTGCTCCAGGAACACACAGTCTGCTCAGAAAGGCGAGATCTCCATTTTAGGTAT 2241
2233 GAGAATTTGCTCCAGGAATACCCATGATGATGATGATGATGATGATGATGATGATGATG 2292
2242 GATCATTTGCCCAAGGATAACCCGATGATGATGATGATGATGATGATGATGATGATGATG 2301
2293 TGTGCTCTGACGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2352
2302 TGTGCTCTGACGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2361
2353 GAAAAATATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2412
2362 GAAAAATATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2421
2413 GCGAAGGAGATTTATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2472
2422 GCGAAGGAGATTTATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2481
2473 TCTCTTACAAACCCAGGAAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2532
2482 TCTCTTACAACTCAGAAAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2541
2533 TCTCAGAGATGCTCAGAGTCACTTAAACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2592
2542 TCAAGAGATGCTCAGAGTCACTTAAACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2601
2593 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652
2602 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2661
2653 TCTGAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2712
2662 TCTGAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2721
2713 AAGCTCTCAACCTGCTCAACACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2772
2722 AAGCTCTCAACCTGCTCAACACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2781
2773 AGTCTAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2832
2782 AGTCTAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2841
2833 GCGAGCTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2892
2842 GCGAGCTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2901
2893 GTGCTCTCTTCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2952
2902 GTGCTCTCTTCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2961
2953 GAAAAATTTTTCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3012
2962 GAAAAATTTTTCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3021
3013 TGTACTGATCTCCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3072
3022 TGTACTGATCTCCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3081
3073 GTGAAGATGCTTCCAGGCTCTTACAGGCTCTTACAGGCTCTTACAGGCTCTTACAGGCTCTT 3132
3082 GTGAAGATGCTTCCAGGCTCTTACAGGCTCTTACAGGCTCTTACAGGCTCTTACAGGCTCTT 3141
3133 AGCATGCTGTAGAGGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3192
3142 AGCATGCTGTAGAGGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3201
3193 TGCACGCTCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3252
3202 TGCACGCTCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3261
3253 ACAGGCAAGCAGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3312

Db	549	CATTGAAGCAGCCAAATAGAAATATGAGACTTCGCTTCAACCAATTTTTCCTACAGAAATG	608
Qy	543	TAGTTCGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC	602
Db	609	TAGCTGGGACCAATTTATATGTTTATGATGGGACTCAATCTACGCACCTCTGATTTGCTGC	668
Qy	603	ATTATAGTGGCTCATTTGTTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTC	662
Db	669	CTTTAGTGGCTCATTTGTTTCTGAAAGAGATGGCAATGAGACGGCTCTGAGGTCACGTGT	728
Qy	663	CACATCAGGTTATGCTTGTCTGCAATTTTATAGTGATGCTGTATATAATTTCACTGGATT	722
Db	729	CACCTCAGGTTATGCACTGCTGCAATTTTTCAGTGATGCTTATATATCTGACTGGAAT	788
Qy	723	TAATATTACTTACAGTTTGTATATGTCTCAATAATACTCTCAGCCGAGGAGAGTGTAA	782
Db	789	TAATATCACTTACAATTTTGACATGTGTCCGAATAATTCCTCAGGCCGAGGAGGTGTAA	848
Qy	783	GATCAGTAATAGCAGCGAAACTGTTCAATGTGAATGTTCTGAAAACTGGAAGGTGAAGC	842
Db	849	GAGCAGTAAACAGCAGACGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAGGGAGTTC	908
Qy	843	ATGTGACATTCCTCACTGTACAGACCAACTGTGGTTTTCTCTCATCGAGGCAATCTGCAATTC	902
Db	909	GTGTGACATTCCTCACTGTACAGACCAACTGTGGTTTTCTCTCACCAGGAGCATCTGTAATGC	968
Qy	903	AAGTGATGTACAGAGGATGCTCTGCTTCTCAGACTGCGAGGTCCTGGAATGTTCAGTTCC	962
Db	969	AAGCGATACAGAGGGTGCTCTGCTTCTCTACTGGCAGGTCCTGGAATGTTCAATTC	1028
Qy	963	TGTACCAAGCTAAACCAAGTCATTTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCCAG	1022
Db	1029	TGTGCCAGCTAAACCAAGTCATTTTGGACTCGAGAGGAATATTTCTGATTTAAAGCTTCCCAG	1088
Qy	1023	AGCATCTCATAAAGCTGTGTGTAATGGAACATATATGTGGTTGTGGAGGATATATGTT	1082
Db	1089	AGCCTCTCATAAAGCTGTGTGTAATGGAACATATATGTGGTTGTGGCGGATATATGTT	1148
Qy	1083	CAACCACTCAGATTTATAACATGGTTCTAGCGTATGACCTTTCCTTCTAGGAGGTGGCTCC	1142
Db	1149	CAACCATTCAGATTTACAGCATGGTTTACGATGACCTGACCTCTAGGGAATGGCTTCC	1208
Qy	1143	ACTAAACGGTTCGTGAAACAAATGTGTTTATGATATGTCATTCTTTGGCATTATACAA	1202
Db	1209	ACTAAACCATTCCTGTGAACAGTGTGTTGTGAAGATATGTCATTCTTTGGCATTATACAA	1268
Qy	1203	GGATAAAATTTACATGTATGAGGAAAAATGTATCCAACTGGGAATGTGACCAATCAGTT	1262
Db	1269	GGATATAAATCTACATGTATGGAGAAAAATTTGATTTCAACAGGGAACTGACCAATCAGCT	1328
Qy	1263	GAGAGTTTTTCAATTCATAATAGTCAATGGTGTGTTTGACCCCTTAAGCCAAAGAGCA	1322
Db	1329	GAGAGTATTTTCATATTCATAATGAATCATGGGTATTTGTTAACTCCGAAAGCTAAGGATCA	1388
Qy	1323	GTATGCAGTGTTGGGCACTCTGCACACATTTGTACACTGAAGAAATGGCCGAGTGTGCAT	1382
Db	1389	GTATGCAGTGTTGGCACTCAGCACACATTTGTACACTGGATCTGGCCGTGTGTCAT	1448
Qy	1383	GCTGGTCATCTTTGGTCACCTGCCCTCTCTATGGATATATAAGCAATGTGCAGAAATATGA	1442
Db	1449	GTTGGTCATCTTCGGTCAATTCGCCCACTCTATGGATATAAAGCGTTGTGCAGGAATATGA	1508
Qy	1443	TTTGGATGAAGAACACATGGAGTATATTACACACCCAGGTCGCCCTTGTGCAAGGGGTTA	1502
Db	1509	CTTGGAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGGGGTTA	1568
Qy	1503	CGGCCATAGCAGTGTATTACGACCATAGAACCCAGGGCCCTATAGCTTTCATGGTGGCTACAA	1562
Db	1569	TGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTTCATGTTGGCTACAA	1628
Qy	1563	GGCTTTCAGTGCCAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC	1622
Db	1629	GGCTTTCAAGCGCAACAAATACCGGCTTGAGATGAGCTCTTACAGATACAGATGGGATAC	1688

Qy	1623	CGAGTGTGGACCAATCTCTTTAAGGACAGCCGATTTTTCCGGTACTTTCGACACACAGCTGTGAT	1680
Db	1689	TCAGATGTGGACCAATCTCTTTAAGGACAGCCGATTTTTCCGGTACTTTCGACACACAGCTGTGAT	1748
Qy	1683	AGTGAGTGGAAACCATCTGGTGTTTGGGGGAAACACACAATGACACACTCTATGAGCCA	1742
Db	1749	AGTGAGTGGAAACCAATCTGGTGTTTGGAGGGAACACACAATGACACACTCTATGAGCCA	1808
Qy	1743	TGGCGGCAATGTTCTCTTCAGATTTCATGGGCTATGACATTCCTGTGACCGCTGGTCT	1802
Db	1809	CGGTGCGAAATGTTCTCTCGGACTTCATGGGCTTATGACATTCCTGTGACCGATGGTCT	1868
Qy	1803	AGTGCTTCCAGACCTGATCTCCACCATGATGTCACAGATTGGCCATTTCAGCAGTCTTT	1862
Db	1869	AGTGCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTGGCCATTTCAGCAGTCTTT	1928
Qy	1863	ACAAACAGCACCATGATGTGTTCGGTGGTTTCAATAGTCTCTCTCCTCAGCGACATCTCT	1922
Db	1929	GTAAACAGCACCATGATGTGTTCGGGCTTCAACAGCCTCTCTCCTCAGTGACGTCTTT	1988
Qy	1923	GGTATTCACCTCGGACAGTGTGATGGCATCGGAGTGAAGCGCTGTGTTTACGACGAGG	1982
Db	1989	GGTCTTTACCTCGGAGCAGTGGCATGCACACCCGAGTGAAGCTGCTTTGTGTGCAGCAGG	2048
Qy	1983	ACCTGCTATTCCGTTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCTGGGCGCTGGC	2042
Db	2049	ACCTGGTATCCGGTGCTGTGGGACACACAGTCTGTCTCGATGTACCTCTCGGAGTTGGC	2108
Qy	2043	AACGTATGAACAGAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCA	2102
Db	2109	AACGTAGAAACAAAGCAGAAAAGTTAAATCAGAGTGTTTTTCTTAAAGAAACCTCTGACCA	2168
Qy	2103	TGACAGATGTGACGACACACAGATTGTTACAGCTGTACAGCCAAACCAATGACCTGCCA	2162
Db	2169	TGACAGATGTGACGACACACAGATTGTTACAGCTGCACAGCCAAATACCAATGACCTGCCA	2228
Qy	2163	CTGGTGCATGACCAATTGTGTCCCGAGAACACACAGCTGTCTAGAGGCGCAGATCTCCAT	2222
Db	2229	CTGGTGCATGATCACTGTGTCTCTGTGAACACACAGCTGCACAGAGGCGCAGATCTCCAT	2288
Qy	2223	TTTTAGGTATGAGNAATTTGCCCCAGAGTAACCCCATGTACTACTGTATAACAGAGACCCAG	2282
Db	2289	TGCCAAGTATGAGAGTTTGCCCCAAAGGATAACCCCATGTACTACTGTCAATTAAGAAACCCAG	2348
Qy	2283	CTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCGGGAATCAGAGTGCAT	2342
Db	2349	CTGCAGAGCTGTGCCCTAGACCAGAACTGCCAGTGGAGCCCGGGAATCAAGAGTGCAT	2408
Qy	2343	TGCCCTGCCCGAAAATATCTGTGGCAATTGGCTGGCAATTTGGTTGGAACCTCATGTTTGAA	2402
Db	2409	CGCCCTGCCCGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAACCTCGTGTCTGAA	2468
Qy	2403	AATTACTACTGCCAAGGAGAAATATGACAAATGCTTAAATGTTCTGTAGGAACACCAATGC	2452
Db	2469	AATCACTACTGCTAAGGAGAAATATGACAAATGCTTAAATGTTCTGTAGGAACACCAATGC	2528
Qy	2463	CTTTTGGCTTCTCTTAAACCCAGAGAGGTAAGAAATTTGCTCCTTAAGCAGCTGCCAAT	2522
Db	2529	CTTTTGGCTTCTCCTCACAATCCCAAGAGAGGTGGAGTTTGTCTCTTAAGCAGCTTCGATT	2588
Qy	2523	AATGAGTCACTCAGAGCATGTCCAGCTCAACCTTAACCCCATGGGTGCGCCTTCGAA	2582
Db	2589	AATGCAATCATCTCAAGTAATGTCCAGCTCACTCTGACTCCATGGGTGGTCTTCGGAA	2648
Qy	2583	GATCAATGTCTTACTGGTGCTGGGAGATATGTCCTCCCAATTTTCAAAATAGTTTACTACA	2642
Db	2649	GATCAATGTCTTACTGGTGCTGGGAGATATGTCCTCCCAATTTTCAAAATAGTTTACTACA	2708
Qy	2643	GTGAGTCCGCTCTGAGCCAGTGAATGCTGGATTCGTGGGAATTTTATCAGAACCCAGTAC	2702
Db	2709	GTGAGTCCATCTGAGCCAGTGAATGCTGGCTCTGTGGGATCTGTGAGAGCTGTGATGAC	2768

QY 2161 CACTGGTGAATGACCAATGTGTGCTCCAGAAACCAACAGCTGCTCAGAAAGCCAGATCTCC 2220
DB 2155 CACTGGTGAATGATCACTGTGTGCTCCAGAAACCAACAGCTGCTCAGAAAGCCAGATCTCC 2214
QY 2221 ATTTTATGATGAGAAATGCTCCCAAGGATTAACCCCATGTACTACTGTAACAAAGAAC 2280
DB 2215 ATTTCAAGTATGACAAATGCTCCCAAGGATTAACCCCATGTACTACTGTAACAAAGAAC 2274
QY 2281 AGCTGAGAGCTGTGCTCCAGAAACCAAGCTGCTGAGGAGCCCGGAAATCAGAGTGC 2340
DB 2275 AGCTGAGAGCTGTGCTCCAGAAACCAAGCTGCTGAGGAGCCCGGAAATCAGAGTGC 2334
QY 2341 ATTGCCCTGCCGAAATATCTGTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 2400
DB 2335 ATGCTCTGCTGCAATATCTGTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 2394
QY 2401 AAAATTTACTGCTCCAGAGGAAATATGACAAATGCTGCAATGCTGCAATGCTGCAATGCT 2460
DB 2395 AAAATTTACTGCTCCAGAGGAAATATGACAAATGCTGCAATGCTGCAATGCTGCAATGCT 2454
QY 2461 GCGCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTCTTAAAGCAGCTGCGA 2520
DB 2455 GCGCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTCTTAAAGCAGCTTGA 2514
QY 2521 ATAATGCACTATCTCAGAGATGCTCAGAGCTCACTTAAACCCATGGGTGCGCTTCGG 2580
DB 2515 TTAATGCACTATCTCAGAGATGCTCAGAGCTCACTTAAACCCATGGGTGCGCTTCGG 2574
QY 2581 AAGATCAATGCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
DB 2575 AAGATCAATGCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2634
QY 2641 CAGTGATGCTCTGAGCCAGTGAATGCTGAGTATCTGAGTATCTGAGTATCTGAGTATCT 2700
DB 2635 CAGTGATGCTCTGAGCCAGTGAATGCTGAGTATCTGAGTATCTGAGTATCTGAGTATCT 2694
QY 2701 ACTCGGGGACTGAGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
DB 2695 ACTCGGGGACTGAGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2754
QY 2761 CTGCAAAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
DB 2755 CTGCAAAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2814
QY 2821 GATTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
DB 2815 GAGTGCACTAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2874
QY 2881 TCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
DB 2875 TCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2934
QY 2941 ACTGCCCCCTGAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
DB 2935 AGCTGCCCCCTGAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2994
QY 3001 GGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
DB 2995 GGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3054
QY 3061 TATAAGGACCAAGTGAAGTGCCTTCCAGGCTTCTACAGGCTTCTACAGGCTTCTACAGGCTTCT 3120
DB 3055 TATAAGGACCAAGTGAAGTGCCTTCCAGGCTTCTACAGGCTTCTACAGGCTTCTACAGGCTTCT 3114
QY 3121 CTGCTCAATTTCCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
DB 3115 CTGCTCAATTTCCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3174
QY 3181 CCAGCTTCCCAATGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
DB 3175 CCAGCTTCCCAATGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3234

QY 3241 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
DB 3235 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3294
QY 3301 ACCAATGAGGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
DB 3295 ACTAATGAGGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3354
QY 3361 AACACGGGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
DB 3355 AACACGGGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3414
QY 3421 GAGTGAATAATCGATACCAAGGAAACCTCTCTCAGAGGAAACATGTTTATTTATATCTTCTT 3480
DB 3415 GAGTGAATAATCGATACCAAGGAAACCTCTCTCAGAGGAAACATGTTTATTTATATCCCTTCTC 3474
QY 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATC 3540
DB 3475 ATTGACTATCAGTTCACTTTAGTCTATCCAGAGAGATGATCGCTATTATATACAGCTATC 3534
QY 3541 AATTTTGTGCTACTCTCTGACGAAACAAACAGGAGTTTGGACATGTTTCAATGCTCTC 3600
DB 3535 AATTTTGTGCTACTCTCTGACGAAACAAACAGGAGTTTGGACATGTTTCAATGCTCTC 3594
QY 3601 AAGAAATTTCACTCAACATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3660
DB 3595 AAAAACTTCACTCAACATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3654
QY 3661 GAAGAGATGCTGCTGCTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
DB 3655 GAAGAAATGCTGCTGCTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3714
QY 3721 GAGAAATTTGATTTTCGCAACCAACCAATTTCTTTTATGCTGCTGCTGCTGCTGCTGCTGCT 3780
DB 3715 GAGAAATTTGATTTTCGCAACCAACCAATTTCTTTTATGCTGCTGCTGCTGCTGCTGCTGCT 3774
QY 3781 ACCTGCCCCATCAAAATTCAGGT 3803
DB 3775 ACCTGCCCCATCAAAATTCAGAT 3797

RESULT 8

AB038387
LOCUS AB038387 8739 bp mRNA linear ROD 18-JAN-2001
DEFINITION Rattus norvegicus atrn mRNA for attractin, complete cds.
ACCESSION AB038387
VERSION AB038387.1 GI:12275307
KEYWORDS attractin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,
Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T.
and Serikawa, T.
TITLE Attractin/mahogany/zitter plays a critical role in myelination of
the central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
MEDLINE 21143347
PUBMED 11209055
REFERENCE 2 (bases 1 to 8739)
AUTHORS Kuramoto, T., Serikawa, T. and Ushijima, T.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division, Tsukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp,
Tel:81-33542-2511(ex.4521), Fax:81-33565-1753)
FEATURES
source
1..8739
/organism="Rattus norvegicus"
/mol_type="mRNA"

```

/strain="WTC"
/db_xref="taxon:10116"
1. .8739
/gene="atrin"
61. .4359
/cds
/gene="atrin"
/codon_start=1
/product="attractin"
/protein_id="BAB21017.1"
/db_xref="GI:12275308"
/translation="MVAIAAAAEATEARLGRYTTAATAAPAGKEROHRPCAATGAWRP
WPRAGLCPLVLSALSPPLPLPLPLLSLPLPREAAEAAVAASGSAARAK
ECDPVCNGRCNGPCGQCVCPTGWCQEQHCGGRFLRTGSGSFVDTGPNKYTK
CTWLEEQPNKIMLRPNHFAATECSMDHLVYDGDLSIYAPLIAFSLIVPERDNET
APEVTVSGVALLHFFSDAANLUGFNITTFDMCPNCSGRGECSSNSSSVBEC
SENWKGSCDIPHCITDNGCPHRCINASDTGSCFPHWQPGCSIPVANOSWTR
EYVDLPLPASHKAVNGNIMWYGVGNFNHSDYSMLAYDLASREWLSNHSVSV
VYRGSLALHDKIKIYVNGKIDTGNVNLRLVFIHNSWVLLTPKADQVAVVGH
SAHTVTLSSGRVLMVIFGHCPLYGYSVQVEYDLEKNTWSILQTCQALVQGVGHSS
VYDRTKALVYHGVYKAFSAKXYELADLVYHVDTOMWTILKDSRPFRLVHTAVVS
GTMVLVFGNTDNTSMHGAKCPSSDFMAYDIACDRMSVLPRPDLHVDVNRFGHSVL
HNSWYVFGGNSULLSDVLFTSEQDQAHRSERACVAAGPIRCLWDVTSRCSWNE
LATEQAEKLSKESKLTLDHRCDDQHTDCTANTNDCHNDHCVPNHSCWTEG
QISATDNCDFPNMPCYCNKTSRSCALDQNCQWEPNQBECIALPENICGHWLV
GNSCLKITAKENDYNAKLSRHNENAFSLTSOKVEFVLKOLRLMQSSQSTKLT
TPWGLRTKIYVSWCWDMSPTNSLLQWMPSPSDAGFCGLSEPTSGRLKAATCIN
PLAGSVERPANYSKACORPCALRTAGCCTSSSRMCMSNMKQCVDSNAVYAGFP
FGQWENYTMSCPEPNCGYCTCSHCLQPGCWCTDPBNTGKKGCIEGSYKPVKM
FSHASTGNVYPPQLNSMCLSDRYNWSFIHCPACQCNHSGKINSICKECEDLT
KGHCETISYGPDTNGKQCPCKNGHASLNTNTGKFCFTTGKVGEECOLCEVE
NRYQCNPLKGTCTYTLIDYQFTSLQEDDRYTAINFVATDEQRNDLDMFNASK
NFNLIWATSPAGTGTGVEVPVSKTNIKEYKDSFSEKFPFRNHPNITFFVYVSN
PTWPKIQIASPSHNSMVDLVQFVTFPFCFLSLLVAVVKIKQSCWASRRBOLL
REMOMASRPASVNALETDEEPDLIGSITVPKPIALEPCFGKNAALVAVFREL
PRGLGGPFPFGQSGLIAVASALVDLSQMPPIVYKESGAVNRKQPPAQPGTCT"
ORIGIN
Query Match      81.6%; Score 3114.8; DB 10; Length 8739;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 3371; Conservative 0; Mismatches 427; Indels 0; Gaps 0;
QY 3 GGTGCCCGCGCGCGGCACTGAGGCAAGCGCTGAGGAGGAGGACGGCGGCGAGCGGAGC 62
DB 72 GCGCGCGCGCGCGGCGGACTGAGCGCGGCTGAGGCGGCTACACGACGGCGGACGAGC 131
QY 63 GCTCGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122
DB 132 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 191
QY 123 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
DB 192 GTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 251
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
DB 252 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
QY 243 CGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302
DB 312 CGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 371
QY 303 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 362
DB 372 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 431
QY 363 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 422
DB 432 CACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 491
QY 423 TTCTGCGGTTGTGACAGATGGAATTAATAATAATAATAATAATAATAATAATAATAATA 482
DB 492 TTCTGCGGTTGTGACAGATGGAATTAATAATAATAATAATAATAATAATAATAATAATA 551
```

```

QY 483 CATTGAAGACAGCAGCAATAGAAATAATGAGACTTCGTTTCAATCATTTTGTCTACAGAGTG 542
DB 552 CATTGAAGACAGCAGCAATAGAAATAATGAGACTTCGTTTCAATCATTTTGTCTACAGAGTG 511
QY 543 TAGTTGGGACCAATTTATATATGTTTATGATGGGACCTCAATTTATGACCCCTAGTTGCTGC 602
DB 612 TAGTGGGACCAATTTATATATGTTTATGATGGGACCTCAATTTATGACCCCTAGTTGCTGC 671
QY 603 ATTAGTGGCCTCATTTGTTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGC 662
DB 672 CTTTATGTCCTCATTTGTTCTCTGAGAGATGGCAATGAGAGCGGCCCTGAAAGTCACTGT 731
QY 663 CACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
DB 732 CACTTCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
QY 723 TAATATCTTACAGTTTGTATGATGTCCTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
DB 792 TAATATCACTTCAATTTTGTGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
QY 783 GATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACTGGAAGAGTGGAAGC 842
DB 852 GAGCAGTAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 911
QY 843 ATGTGACATTCCTCACTGTACAGACAACCTGTGGTCTTCTCATCGAGGATCTGCAATTC 902
DB 912 CTGTGACATTCCTCACTGTACAGACAACCTGTGGTCTTCTCATCGAGGATCTGTAACGC 971
QY 903 AAGTGATGTCAGAGGATGCTGCTGCTTCTCAGACTGGCAGGTCCTGGAATGTTCACTTCC 962
DB 972 AAGTGACACAGAGGATGCTGCTGCTTCTCAGACTGGCAGGTCCTGGAATGTTCACTTCC 1031
QY 963 TGTACAGAGTAAACAGCTCACTTTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCCAG 1022
DB 1032 TGTGCGGCGCTAACCGACTCTTTTGGACTCGAGAGGAATATTTCTGACTTAAAGCTCCCCAG 1091
QY 1023 AGCATCTATAAGCTGTGCTGCTCAATGGAACATTATGTCGGTGTGTTGGAGGATATATGTT 1082
DB 1092 AGCCTCTCTATAAGCTGTGCTCAATGGAATATATGTCGGTGTGTTGGCGGATATATGTT 1151
QY 1083 CAACCACTCAGATTATAACAATGTTCTAGCGTATGACCTTCTTCTAGGAGTGGCTTCC 1142
DB 1152 CAACCACTCAGATTATAACAATGTTCTAGCGTATGACCTTCTTCTAGGAGTGGCTTCC 1211
QY 1143 ACTAAACCGTCTGTGAAACAATGTTGTTAGATATGTTGTTGCTGCTGCTGCTGCTGCTGCT 1202
DB 1212 ACTAAACCACTTCTGTGAACAGTGTGTTGTAAGATATGTTCTTCTGCTGCTGCTGCTGCT 1271
QY 1203 GGATAAAATTTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1262
DB 1272 GGATAAAATTTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1331
QY 1263 GAGAGTTTTTTCACATTCATAATGAGTCACTGGGTTGTTGACCCCTTAAGCAAGAGCA 1322
DB 1332 GAGAGTTTTTTCATAATTCATAATGAGTCACTGGGTTGTTGACCCCTTAAGCAATCA 1391
QY 1323 GTATGAGTGGTGGGACCTCTGACACATTGTTTACACTCAAGAAATGGCGGAGTGGTCA 1382
DB 1392 GTATGAGTGGTGGGACCTCTGACACATTGTTTACACTCACTGCTGCTGCTGCTGCTGCTGCT 1451
QY 1383 GCTGGTCACTTCTGGTCACTGCTGCTCTCTATGATATATAAGCAATGTGACGAATATGA 1442
DB 1452 GTTGGTCACTTCTGGTCACTTCTGGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
QY 1443 TTTGGATAAGAACACATGAGTATATTAACACCCAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
DB 1512 CTTTGGAAAGAACACATGAGTATATTAACAACTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1571
QY 1503 CGGCCATAGCAGTGTGTTTACGACCATGAGACCGGCCCTATACGTTTCACTGCTGCTGCTGCTGCT 1562
DB 1572 TGGCCACAGCAGTGTGTTTATGATACAGACCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1631
QY 1563 GGCTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
```

Db 1632 GGGCTTTTCAGCGCCAAACAGTACCGGCTTGCGATGACCTCTACAGATACATGTGGACAC 1691
QY 1623 CCAGATGTGGACCATCTTTAAAGACAGCGGATTTTCGGTTACTTTCGACACAGCTGTGAT 1682
Db 1692 TCAGATGTGGACCATCTTTAAAGACAGCGGATTTTCGGTTACTTTCGATACAGCTGTGAT 1751
QY 1683 AFTGAGTGAACCATCTCTGGTGTGTTGGGGGAAACACACACAATGACACATCTTATGAGCCA 1742
Db 1752 AFTGAGTGAACCATCTCTGGTGTGTTGGAGGGAACACACACAATGACACATCTTATGAGCCA 1811
QY 1743 TGGCGCCAAATGCTTCTCTCAGATTTCAATGCGCTTATGACATGCTGTGACCGCTGTGC 1802
Db 1812 CGGTGCGAAATGCTTCTCTCAGATTTCAATGCGCTTATGACATGCTGTGACCGCTGTGC 1871
QY 1803 AFTGCTTCCAGACCTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGACAGTCTT 1862
Db 1872 AFTTCTCTCCAGACCTGTAGCTCCATCAGATGTCAACCGATTTGGCCACTCAGACAGTCTT 1931
QY 1863 ACACAACAGACCATGTATGTGTTTCGGTGGTTTTCATATGATCTCTCTCAGCGACATCTT 1922
Db 1932 GCACAACAGACCATGTATGTGTTTCGGGCGCTTCAACAGCGCTCTCTCAGTGAAGCTT 1991
QY 1923 GGTATTTCACTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTGTTAGCAGCAGG 1982
Db 1992 AGTCTTCACTCGGAGCAGTGTATGTCACACCGCAGCGAGCTGCTGTGTGGACGAGG 2051
QY 1983 ACTCGATTTCCGTGTGTGTGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGCG 2042
Db 2052 ACTCGCATCCGCTGTCTGTGGGACACACAGTCGTCTCGATGTACTCTCTGGAGTTGGC 2111
QY 2043 AACTGATGAACAGAGAAAGTTAAATTCAGATGTTTTCGAAAGATCTCTTGACCA 2102
Db 2112 AACTGAAGAACAGAGAAAGTTAAATTCAGATGTTTTCGAAAGATCTCTTGACCA 2171
QY 2103 TGACAGATGTGACAGCACACAGATTTGTACAGCTGTACAGCCCAACACCAATGACTGCCA 2162
Db 2172 TGACAGATGTGACAGCACACAGATTTGTACAGCTGTACAGCCCAATACCAATGACTGCCA 2231
QY 2163 CTGGTGCAATGACATTTGTGTCGCCAGGAAACACAGCTGTCTCAGAGCCAGATCTCGAT 2222
Db 2232 CTGGTGCAATGACATTTGTGTCGCCAGGAAACACAGCTGTCTCAGAGCCAGATCTCGAT 2291
QY 2223 TTTTAGTATGAGATTTGCCCAAGCATTAACCCATGTAAGTGTAAACAGAGACAG 2282
Db 2292 TGCCAGTATGACATTTGCCCAAGCATTAACCCATGTAAGTGTAAACAGAGACAG 2351
QY 2283 CTGAGGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGGAAATCAGGAGTGCAT 2342
Db 2352 CTGAGGAGCTGTGCCCTGGATCAGAACTGGCCAGTGGGAACTCGAAATCAGGAGTGCAT 2411
QY 2343 TGCCCTGCGCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTGA 2402
Db 2412 CGCCCTGCGCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTGA 2471
QY 2403 AATTACTACTGCCAGGAGAAATATGACAATGCTAAATTTGTTCTGTAGGAAACCAATGC 2462
Db 2472 AATCACTACTGTAAGGAAATTTATGACAATGCTAAATTTGTTCTGTAGGAAACCAATGC 2531
QY 2463 CTTTTTGGCTTCTCTTACAAACAGAGAGGTAGAAATTTGCTCTTAAGCGCTGCGAAT 2522
Db 2532 CTTTTTGGCTTCTCTTACAAACAGAGAGGTAGAAATTTGCTCTTAAGCGCTTGCAT 2591
QY 2523 AATGAGTCACTCTCAGAGCATGTCCAGCTCACTCTCAACCCATGCGGTTCGCGCTTCGAA 2582
Db 2592 AATGCAATCACTCAGAGCATGTCCAGCTCACTCTCAACCCATGCGGTTCGCTTCGAA 2651
QY 2583 GATCAATGTCTCTACTGGTCTGGAGAGATATGTCCCATTTACAAATAGTTTACTACA 2642
Db 2652 GATCAATGTCTCTACTGGTCTGGAGAGATATGTCCCATTTACAAATAGTTTACTACA 2711
QY 2643 GTGATGCGCTCTGAGCCAGTGTGCTGGATTTCTGTGGATTTTATCAGAACCCAGTAC 2702

Db 2712 GTGGATGCCATCTGAGCCAGTGTATCCGGCTTCTGTGGATCTTGTGAGCCTAGTAC 2771
QY 2703 TCGGGACTGAAGGCTGCAACCTGCATCAACCCACCTCAATGGTAGTGTGTGTAAGGCC 2762
Db 2772 TCGGGACTGAAGGCTGCAACCTGCATCAACCCACCTCAATGGTAGTGTGTGTAAGGCC 2831
QY 2763 TGCAAACACAGTGTCTAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCATGTGGAGA 2822
Db 2832 TGCAAACACAGTGTCTAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCATGTGGAGA 2891
QY 2823 TTCGACAGCGGACAGTCTTGAGTGTATGTGTCAGACACATGAAGCAGTGTGAGCTC 2882
Db 2892 GTGCACAGCAGCAGCTTCCGAATGCAATGTGTGACAGCAATGAAGCAGTGTGAGTAC 2951
QY 2883 CAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATACCATGAGCAC 2942
Db 2952 CAATGCTATGTAGCTCTCTTCTTTCGGCCAGTGTATGGAATGTTATACCATGAGCAC 3011
QY 2943 CTGCCCCCTGAAAATTTGTTTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGG 3002
Db 3012 CTGCCCCCTGAAAATTTGCTTGGCTACTGTACTGTAGTCAATGCTTGGAGCAGCAGG 3071
QY 3003 CTGTGGCTGTGTACTGTATCCAGCAATACTGTGCAAGGAAATGATAGAGGTTCTCTA 3062
Db 3072 CTGTGGTGTGTACTGTATCTTACATGCTTGGAGGAAATGATGTTGAGGAGCTA 3131
QY 3063 TAAAGCACAGTGAAGATGCTTTCGACGCCCTACAGGAAATTTCTATCCACAGCCCT 3122
Db 3132 TAAAGCACCGTGAAGATGCTTTCATGCTCTACAGGAAATTTGTTACCCACAGCCCT 3191
QY 3123 GCTCAATTTCCAGCATGTGTAGAGGACAGATACATGCTGCTTTTCACTACCTGCTC 3182
Db 3192 TCTGAATCTCAGCATGTGTGTAGAGGACAGATACATGCTGCTTTTATCTACCTGCTC 3251
QY 3183 AGTTTCCCAATCAACCGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGA 3242
Db 3252 AGTTTCCCAATGCAATGAGACACAGCAAGTGCATCAACAGAGCATCTGTGAGAAGTGA 3311
QY 3243 GAACCTGACACAGGCAAGCATCTGGAGACCTGCATATCTGCTTCTACGCTGATCCAC 3302
Db 3312 GGACCTGACACAGGCAAGCATCTGGAGACCTGCATATCTGCTTCTATGCTGACCCGAC 3371
QY 3303 CAATGAGGAAATGTCAAGTGCATGCAAGTGCATGCAATGGGACCGCTCTGTGCAACACCA 3362
Db 3372 TAATGAGGCAATGTCAAGTGCATGCAAGTGCATGCAAGGACCGCATCACTGTGCAACCA 3431
QY 3363 CACGGCAAGTGTCTGTGACCAACCAAGGCGCTCAAGGGGAGAGTGCAGCTATGTA 3422
Db 3432 CACTGCAAGTGTCTGTGTACCAACCAAGGCGCTCAAGGGAGAGTGCAGCTATGTA 3491
QY 3423 GGTAGAAAATCGATCAACAGGAAACCTCTCAGAGAAACATGTTATATCTTCTTAT 3482
Db 3492 GGTAGAAAATCGATCAACAGGAAACCTCTCAGAGAAACATGTTATATCTTCTTAT 3551
QY 3483 TGAATATCAGTTTCACTTTTAGTCTTATCCAGAGATGATCGCTATTAACAGCTATCAA 3542
Db 3552 TGAATATCAGTTTCACTTTTAGCTGTCCAGGAGACAGCCGCTATTAACAGCTATCAA 3611
QY 3543 TTTTGTGGTACTCTCTGACGAAACAGGATTTGGACATGTTTCAATGCTTCTTCA 3602
Db 3612 TTTTGTGGTACTCTCTGATGAAACAGGATTTGGACATGTTTCAATGCTTCTTCA 3671
QY 3603 GAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGA 3662
Db 3672 AAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGAGA 3731
QY 3663 AGAGATGCTGTGTTTCAAAACCAACATTAAGAGATCAAAAGATGTTTCTTAATGA 3722
Db 3732 AGAGTGTGCTGTGTTTCAAAACCAACATCAAGGAATACAAAGATGCTTCTTAATGA 3791
QY 3723 GAAGTTTGTATTTTCCGCAACCCCAATATCACTTTCTTTTGTATGTCAGTAATTTTCA 3782
Db 3792 GAATTTGTATTTTCCGCAACCTCAACATTAATTTCTTTTGTATGTCAGTAATTTTCA 3851


```

3783 CTGCCCCATCAAAATTC 3800
3852 TTGCCCCATCAAAATTC 3869

RESULT 9
AB038388 4500 bp mRNA linear ROD 18-JAN-2001
LOCUS Rattus norvegicus Atrn mRNA for attractin, complete cds.
DEFINITION AB038388
ACCESSION AB038388.1 GI:12275311
VERSION attractin.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (sites)
Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,
Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T.
and Serikawa, T.
Attractin/mahogany/zitter plays a critical role in myelination of
the central nervous system
Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
2143347
11209055
2 (bases 1 to 4500)
Kuramoto, T., Serikawa, T. and Ushijima, T.
Direct Submission
Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp,
Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
Location/Qualifiers
1. 4500
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="WTC"
/db_xref="taxon:10116"
1. 4500
/gene="Atrn"
61. 3888
/gene="Atrn"
/codon_start=1
/product="attractin"
/db_xref="GI:12275312"
/protein_id="BAB21018.1"
/translation="MVAAAABAEARLRYTTATAPAPAKWKRQHRPCATCAWPR
WPRAGLCPLRVLSRALSPPLPLPLPLPLPLPLPREAABAAVAAVSGAAAEAK
ECDRPCVNGRCNPQGCQVPTGWGQCQCGRFLTSGSGFVTDGPNKYKTK
CTWLLEGOPNKMRLRFNFHATECSWDHLYVVDGDSIYAPLIAAFSGLIIVERDNET
APEVTVTSYALLHFTSDAAYNLITGFNITNFDMCPNCSGRGCKSNSSTVECEC
SENWKGSCDIPHCITDNCQFPHRGICNASDTRGCGCFPHWQPGCSLPVPANQSFWR
EYSDLKUPRASHKAENVGNIMWVGGFMFNFHSDYSVVLADLASREHLSLHNSVSV
VRYGHSALAKHDKIYMYGGKIDISTGVNTELRVPHIHNSWLLITPRAKQYAVVGH
SAHIVTSLSGRVLMVIFGHCPYGIYISVQYEDLEKNTWLSILOQALVQGGYHSS
VYDRTKALYVHGGYKAFSANKYRLADDLRYHYVDTONMTWIKDSRFRFYLHTAVIVS
GTMVLFVGNTHNDTSMHGAKCFSDFMAYDIACDRWSVLPRLPHLDVNFHGSVIL
HNSWYVFGFNLSLLSDVLAFTSQCDHARSEACVAAGPCIRCLWDTQSSRCTSW
LATEQAELKCEKSKTLDHRCDDQHTDCVSCSTANTNDCHWNCNDHCVFVNSHCTEG
QISIARYKNCNDKPNMYNCKTSCRSALDNCOWEPNEBCEIALPNIIGIWHLV
GNSCLKITAKENYDNALKSCTRNHNAFLASLTSOKKVEFLVKQLRMOSOSTSKLIL
TPWGLKRTNVNSWEDMSPNTSLNLMPEPSDAGFCGILSEPSRGLKATIN
PLNGSVCEPANHSAKQRTPCALRTAGCECTSSSESCWCSNMKQCVDSNAYVAFSP
FQGWVETWMSCPENCQCYCTCSHCLQPCQWCTDPSNTGKGCIEGVYKGPVM
PHASTGNVYPOPLNNSMCLDSRYNMSFHLCPACQCNHSHKCIQSIQCEKEDLT
GHCECTISGFTGDPNGKQPCQKNGHSLCNTNTGKCTCTKGVGESQCLVE
NRYQGNPLKTCYLLIDYQFTLSQEDDRIYTAINFVATPDSQNRDLDMFINASK
NFNLNITWATSPFAGTGTGEEVVPVSKINIKYKDSFSENEKFPDFENHNIITFFVIVS
FTWPIKIQVRVTS"

FEATURES
source
gene
CDS
```


Db 1092 ||||| AGCCTCTCAATAAGCTGAGTCAATGGAATATAATGTTGGTGGTATATGTT 1151
QY 1083 CAACCACTCAGATTATAAATGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTCC 1142
Db 1152 CAACCACTCAGATTACAGATGTTTACGCTATGACCTGGCTTCTAGGGAATGGCTTC 1211
QY 1143 ACTAAACCGTCTGTGAAACAAATGTTGTTTATGATATGCTATCTTTGGCAATATACAA 1202
Db 1212 ACTAAACCAATCTGTGAAACAGTGTGTTTAAATATGCTCATCTTTGGCAATATACAA 1271
QY 1203 GGATAAAATTTACATGATGAGGAAATTTGATCCACTGGGAATGTGACCAATGAGTT 1262
Db 1272 GGATAAAATTTACATGATGAGGAAATTTGATCCACTGGGAATGTGACCAATGAGTT 1331
QY 1263 GAGAGTTTTCATATTCATTAATGATCAATGGTGTGTTGACCCCTAAGGCAAAAGGACA 1322
Db 1332 GAGAGTTTTCATATTCATTAATGATCAATGGTGTGTTGACCCCTAAGGCAAAAGGACA 1391
QY 1323 GATGCAAGTGGTGGCACTCTGCAACATTTGTTACATGTAAGAAATGGCCGAGTGGTCA 1382
Db 1392 GATGCAAGTGGTGGCACTCTGCAACATTTGTTACATGTAAGAAATGGCCGAGTGGTCA 1451
QY 1383 GCTGGTCACTTTGGTCACTGCGCTCTCTATGATATATAGCAATGTGCAAGGAATATGA 1442
Db 1452 GTTGGTCACTTTGGTCACTGCGCTCTCTATGATATATAGTGTGTTGCAAGGAATATGA 1511
QY 1443 TTTGGATAGAACATGAGATATATACACACCAAGGCTGCTTGTGCAAGGGGTTA 1502
Db 1512 CTTGGAAAGAACACATGGAGTATATTAACAACTCAAGGTGCTTGTAGTCAAGGGGTTA 1571
QY 1503 CGGCCATACAGTGTGACACCATAGGACAGGCGCTTATAGTGTGCTAGTGGCTCAA 1562
Db 1572 TGGCCACAGAGTGTATGATCAGGACCAAGGCTCTGATGTTCAATGTTGGCTCAA 1631
QY 1563 GCTTTTCAAGTGCATATAGTACCGCTTGCAGATGATCTTACCGATATGATGTTGATAC 1622
Db 1632 GCTTTTCAAGTGCATATAGTACCGCTTGCAGATGATCTTACCGATATGATGTTGATAC 1691
QY 1623 CCAGATGTTGACCATCTTACGACAGCGGATTTTCGTTACTTGTGACACAGCTGTGAT 1682
Db 1692 TCAGATGTTGACCATCTTACGACAGCGGATTTTCGTTACTTGTGACACAGCTGTGAT 1751
QY 1683 AGTGAAGTGAACCATGCTGTTGTTGGGGGAAACACACACATGACACATCTATGAGCCA 1742
Db 1752 AGTGAAGTGAACCATGCTGTTGTTGGGGGAAACACACACATGACACATCTATGAGCCA 1811
QY 1743 TGGCCCAAAATGCTTCTCTCAGATTTATGCGCTATGACATTTGCTGTGACCGCTGGTC 1802
Db 1812 CGGTGCCAAATGCTTCTCTCAGATTTATGCGCTATGACATTTGCTGTGACCGATGTC 1871
QY 1803 AGTGTCTCCAGACCTGATCTCCACATGATGTCAAAGATTTGCGCATTTACAGAGTCTT 1862
Db 1872 AGTGTCTCCAGACCTGATCTCCACATGATGTCAAAGATTTGCGCATTTACAGAGTCTT 1931
QY 1863 ACACAAAGCAACCATGATGTTGTTGGTGGTTCATAGTGTCTCTCTCAGCGACATCCT 1922
Db 1932 GCACAAAGCAACCATGATGTTGTTGGGCGGTTCACACGCTCTCTCTCAGTGAAGTCTT 1991
QY 1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCGCTGTTTATGACGAGG 1982
Db 1992 AGTCTTACCTCGGAGCAGTGTGATGACACCGGAGGAGTCTGTTGTGGGAGGAGG 2051
QY 1983 ACCTGGTATTCGGTGTGTGGAACACAGAGGTGCTCTCAGTGTATCTGTTGGGCGCTGGC 2042
Db 2052 ACCTGGCATTCGGTGTGTGTGGGACACACAGTGTCTCGATGTACCTCTGGGAGTTGGC 2111
QY 2043 AACTGATGAACAGAGAGAAAGTTAAATCAGATGTTTTCGAAAGAACTCTTGACCA 2102
Db 2112 AACTGAAGAAACAGAGAGAAAGTTAAATCAGATGTTTTCGAAAGAAACCTTTGACCA 2171
QY 2103 TGACAGATGTGACCAAGCACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCA 2162

Db 2172 TGACAGATGTGACCAAGCACAGATTTGTTACAGTGTGACAGCCAAATACCAATGACTGCCA 2231
QY 2163 CTGTGCAATGACCATTTGCTCCCGGAAACACAGCTGTCTCAGAGGCGGAGATCTCCAT 2222
Db 2232 CTGTGCAATGATCATGCTGTCTCCCTGTGAACACAGCTGTGACAGAGGCGGAGATCTCCAT 2291
QY 2223 TTTTGTGATATGAGAAATTTGCCCAAGGATACCCCATGTACTCTGTAAACAAGAACCCAG 2282
Db 2292 TGCCAGATATGACAAATTTGCCCAAGGATACCCCATGTACTCTGTAAACAAGAACCCAG 2351
QY 2283 CTGCAAGAGTGTGCTCCCTGCAACAGAACTGCAAGTGGAGCCCGGAACTCAGAGTGTCA 2342
Db 2352 CTGCAAGAGTGTGCTCCCTGATATCAGAAATTTGCCAGTGGGAACTCTGAAATCAAGAGTGCAT 2411
QY 2343 TGGCCTGCCGAAATATCTGTGCAATTTGGCTGGCATTTGGTTGGAAATCTCATGTTTGA 2402
Db 2412 CGCCTGCCGAAACATCTGTGCAATTTGGCTGGCATTTGGTTGGAAATCTCTGTCTGAA 2471
QY 2403 AATTACTATGCAAGGAGAAATTTATGCAATGCTAAATTTGTTCTGTAGGAAACCAATGC 2462
Db 2472 AATCACTACTGTAAAGGAAATTTATGCAATGCTAAATTTGCTGTAGGAAACCAATGC 2531
QY 2463 CTTTGTGGCTTCTTTACAAACCAAGAGGTGAGAAATTTGCTTAAACAGCTGGAAT 2522
Db 2532 CTTTGTGGCTTCTTTACAAACCAAGAGGTGAGAAATTTGCTTAAACAGCTGGAAT 2591
QY 2523 AATGCAATCATCTCAGAGACAGTCCAAAGCTCACTCTGACTCCATGGGTTGGTCTCGGAA 2582
Db 2592 AATGCAATCATCTCAGAGACAGTCCAAAGCTCACTCTGACTCCATGGGTTGGTCTCGGAA 2651
QY 2583 GATCAATGTGCTACTGCTGTGGGAGATATGCTCCANTCAACAAATAGTTTGTGCTGA 2642
Db 2652 GATCAATGTGCTACTGCTGTGGGAGATATGCTCCANTCAACAAATAGTTTGTGCTGA 2711
QY 2643 GTGGATGCTGCTGAGCCCAAGTATGCTGGATTTCTGTGAAATTTTATCAGAAACCCAGTAC 2702
Db 2712 GTGGATGCTGCTGAGCCCAAGTATGCTGGATTTCTGTGGATTTCTGTGAGCTTAGTAC 2771
QY 2703 TCGGGACTGAGGCTGCAACCTGCTGATCAACCACTCAATGTTGTTGTTGAAAGGCC 2762
Db 2772 TCGGGACTGAGGCTGCAACCTGCTGATCAACCACTCAATGTTGTTGTTGAAAGGCC 2831
QY 2763 TGCAAAACCAAGTGTGAGCAGTGGGAGACACATGCTGCTTGGAGCAGATGTGGAGA 2822
Db 2832 TGCAAAACCAAGTGTGAGCAGTGGGAGACACATGCTGCTTGGAGCAGATGTGGAGA 2891
QY 2823 TTGCAACAGCGGAGCTGTGAGTGTGATGTTGTCAGCAACATGAAGCAGTGTGAGCTC 2882
Db 2892 GTGCAACAGCGGAGCTGTGAGTGTGATGTTGTCAGCAACATGAAGCAGTGTGAGCTC 2951
QY 2883 CAATGCTATGTTGGCTCTCTCCCTTTGGCCAGTGTGATGGAATGTTATGATGAGCAG 2942
Db 2952 CAATGCTATGTTGGCTCTCTCCCTTTGGCCAGTGTGATGGAATGTTATGATGAGCAG 3011
QY 2943 CTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTAGTCAATTTGTTGGAGCAACCCAG 3002
Db 3012 CTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTAGTCAATTTGTTGGAGCAGCCAG 3071
QY 3003 CTGTGCTGTGTGATGATGCTCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTA 3062
Db 3072 CTGTGCTGTGTGATGATGCTCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTA 3131
QY 3063 TAAAGSACCAAGTGAAGATGCTTCCAAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122
Db 3132 TAAAGSACCAAGTGAAGATGCTTCCAAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3191
QY 3123 GCTCAATTTCCAGCATGTTCTTAGAGGACAGAGATACAACTGGTCTTTTCACTGTCC 3182
Db 3192 TCTGAATCCAGCATGTTCTTAGAGGACAGAGATACAACTGGTCTTTTATTCATGTCC 3251
QY 3183 AGCTTGCCAATCAACGGCCACAGTAAATGATCAATCAGAGCATCTCTGAGAGTGTGA 3242
Db 3252 AGCTTGCAGTGCATGACACAGCAAGTGCATCAACAGAGCATCTCTGAGAGTGTGA 3311

```

gene
1..4313
/ gene="Mgca"
/ note="mahogany candidate"
CDS
27..4313
/ gene="Mgca"
/ note="mahogany"
/ codon_start=1
/ product="attractin"
/ protein_id="AAD25372.1"

```

Query March	81.4%;	Score 3109;	DB 10;	Length 4313;	
Best Local Similarity	98.8%;	Pred. No. 0;			
Matches 3377;	Conservative	0;	Mismatches 415;	Indels 9;	Gaps 1;
<hr/>					
3	GGTGGCGCAGCGCGCGCAACTGAGCGCAAGCTGAGGAGGAGGACGGCGCGCAGCGCAGC	62			
35	GGTGGCGGTGGCGGCGTCTGACTGAGCGCGGCTGAGGGGGAGCAGCAGCGGACAGCAGC	94			
63	GCTCGGGGAGGAGCGGGCGGCGGCACTGGGACTGGGACTGGGACCTGGAGGGCTGGAGGCC	122			
95	GCCTGCGGGCAGGAAGGGGAGGCACACGACCTTCACCGCAGCAGGGGGCTTGGAGGCC	154			
123	GGGCGTGGGGCGCGGCGCTCGCGTCCCGCGGCTGCTGTCTCCACCGCTCGGGCCACACGGCT	182			
155	GGGACCGCGCGCCCGGCTGTGTCTACCGGGGTGTGTCTCGGGGGCGTCTGCCCGCGCC	214			
183	GCTGCTGCTGCTGTTGTGTGCTCCGCGCGCGCTGTGTGCTGTGCTGCTGCGCTGTGAGGC	242			
215	GCTGCTCCGCTGCTCTT-----TTGCTGTGCTGTGCTGCTGCCCGCGGAGGC	265			
243	CGAGGCGCGCGGCGCGCGCGGGTGTGCGGGCTCAGCCGACGCGGAGGCCAAGGAATG	302			
266	CGAGCGCGCTTCGGTGTGGCGCGGGGTGTGCGGCTTCGGCGCAGCCCGAGGCCAAGGAATG	325			
303	TGACCGGCGCTGTGTCAAACGGCGGTGTGTGCAACCGCTGGCACCGGCGCAGTGGCTGCGCC	362			
326	TGACCGGCGCTGTGTCAAACGGCGCGCTGTGCAACCGCTGGCACCGGCGCAGTGGCTGCGCC	385			
363	CGCGGCTGGTGGGCGGACATGCCAGCAGTGGCGGGGCGCTTCACAGCTAACTGGATC	422			
386	CACGGGCTGGGTGGGCGAGCAATGCCAGCAGTGGCGGGGCGGCTTCAGACTAACTGSGTC	445			
423	TTCTGGGTTTGTGACAGATGGACCTTGGAAATTTATAATACAAACGAAAGTGACCGTGGCT	482			
446	TTCTGGATTTGTACAGATGGACCTTGGAAATTTATAATTAAGACGAAGTGACATGGCT	505			

483 CATTGAAGGACAGCCAAATAGATAATGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTG 542
Db
506 CATTGAAGGCTATCCAAATGCGAGTCTAAGGTTAAGATTCAACCAATTTTGTCTACAGAAATG 565
QY
543 TAGTTGGGACCAATTTATATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTTGCTGC 602
Db
566 TAGCTGGGACCAATTTATATATGTTTATGATGGGACCTCAATTTACGACACCTCTCAATTTGCTGC 625
QY
603 AATTAGTGGCTCATTTGTTTCTGAGAGAGATGGCAATGAGCTGTCCTTGAAGTTGTTGC 662
Db
626 CTTTAGTGGCTCATTTGTTTCTGAAAGAGATGGCAATGAGCTGTCCTTGAAGTTGTTGC 685
QY
663 CACATCAGGTTATGCTTCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATT 722
Db
686 CACTTCAGGTTATGCTTCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATT 745
QY
723 TAATATTACTACAGTTTATGATGCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATT 782
Db
746 TAATATTACTACAGTTTATGATGCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATT 805
QY
783 GATCAGTAATAGCAGGAACTGTTGAATGTAATGTTCTGAAACTGGAAGTGAAGC 842
Db
806 GAGCAGTAACAGCAGCGCTGTTGAGTGTGAATGTTCTGAAACTGGAAGGAGTGC 865
QY
843 ATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCATCTGCAATTC 902
Db
866 GTGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCATCTGTAATGC 925
QY
903 AAGTGTATGTCAGAGGATGCTCCTGCTTCTCAACACTGGCAGGCTCCTGATGTTCAAGTTCC 962
Db
926 AAGCGATACAGAGGCTGCTCCTGCTTCTCAACACTGGCAGGCTCCTGATGTTCAAGTTCC 985
QY
963 TGTACAGGTAACAGTCAATTTTGGACTGAGAGGAATATTTCTAACTTAAAGTCCCCAG 1022
Db
986 TGTGCGCAGCTAACAGTCTTTTGGACTCGAAGAAATATTTCTGATTTTAAAGTCCCCAG 1045
QY
1023 AGCATCTCATAAAGCTGTGGTCAATGGAACAACTATGTTGGTGTGTTGAGGATATATGTT 1082
Db
1046 AGCTTCTCATAAAGCTGTGGTCAATGGAACAACTATGTTGGTGTGTTGAGGATATATGTT 1105
QY
1083 CAACCACTCAGATTAACATGGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCC 1142
Db
1106 CAACCACTCAGATTAACATGGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCC 1165
QY
1143 ACTRAACGTTTCTGTGAACAACTGTGGTGTGTTAGATATGTTGATGTTGTTGAGGATATATGTT 1202
Db
1166 ACTRAACCACTTCTGTGAACAACTGTGGTGTGTTAGATATGTTGATGTTGTTGAGGATATATGTT 1225
QY
1203 GGATAAAATTTACATGTTATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1262
Db
1226 GGATAAAATTTACATGTTATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1285
QY
1263 GAGAGTTTTCATCATTAATGAGTCAATGGTGTGTTAGATATGTTGATGTTGTTGAGGATATATGTT 1322
Db
1286 GAGAGTTTTCATCATTAATGAGTCAATGGTGTGTTAGATATGTTGATGTTGTTGAGGATATATGTT 1345
QY
1323 GTATCAGTGGTGGGCTCTGACACATTTGTTACACTGAAGAATGGCCGAGTGGTTCAT 1382
Db
1346 GTATCAGTGGTGGGCTCTGACACATTTGTTACACTGGCACTGGCACTGGCCGTGTGGTTCAT 1405
QY
1383 GCTGTGCTATTTGGTCTATGCTCTCTATGATATATAAGCAATGTGCAAGGAATATGA 1442
Db
1406 GTTGTGCTATTTGGTCTATGCTCTCTATGATATATAAGCAATGTGCAAGGAATATGA 1465
QY
1443 TTTGATTAAGAACACATCGAGTATATTAACACCCAGGTCCTTGTGCAAGGGGTTA 1502
Db
1466 CTTGGAAGAAGAACACATCGAGTATATTAACACTACTCAGGTCCTTGTGCAAGGGGTTA 1525
QY
1503 CGCCATAGCAGTGTGTTACGACCAATAGGACAGGCGCCCTATGCTTATGCTGCTACAA 1562
Db
1526 TGGCCACAGTGTGCTTATGATGACAGGACCAAGGCTCTGATGCTTATGCTGCTACAA 1585
QY
1563 GGCCTTCAGTCCCAATAGTACCGCTTGGAGATGATCTCTACCGATATGATGTGGATAC 1622

1586 GGCTTTTCAGCGCCAAACAAATACCGGCTTCAGATGACCTCTACAGATACGATGTGGATAC 1645
QY
1623 CCAGATGTGACCAATTTCTTAAGGACAGCCGATTTTCCGTTACTTGGACACACAGCTGTGAT 1682
Db
1646 TCAGATGTGACCAATTTCTTAAGGACAGCCGATTTTCCGTTACTTGGACACACAGCTGTGAT 1705
QY
1683 AGTCAGTGGAAACCAATGCTGCTGTTTGGGGGAAACACACAAATGACACATCTATGAGCCA 1742
Db
1706 AGTCAGTGGAAACCAATGCTGCTGTTTGGGGGAAACACACAAATGACACATCTATGAGCCA 1765
QY
1743 TGGCGCCAAATGCTTCTTTCAGATTTATGCGCTATGACATTTGCGCTGTGACCGCTGGTC 1802
Db
1766 CGGTGCGCAATGCTTCTTCTCAGACTTCATGGCTTATGACATTTGCTGTGACCGATGGTC 1825
QY
1803 AGTCGTTCCAGACCTGATCTCCACCATGATGCAACAGATTTGGCATTTCAGCAGTCCT 1862
Db
1826 AGTCGTTCCAGACCTGATCTCCATGATGTCACAGATTTGGCATTTCAGCAGTCCT 1885
QY
1863 ACACACAGCACCATGATGTTGCTGCTGCTTCAATAGTCTCTCTCCTCAGCGACATCCT 1922
Db
1886 GTACACAGCACCATGATGTTGCTGCTGCTTCAACAGCTCTCTCCTCAGTGACGCTCT 1945
QY
1923 GGTATTCACCTCGAAGACAGTGTGATGGCATCGGAGTGAAGCGCTGTTTACGACAGG 1982
Db
1946 GGTCTTTACCTCGAAGACAGTGTGATGGCATCGGAGTGAAGCGCTGTTTACGACAGG 2005
QY
1983 ACCTGCTATTCGGTGTGTGTGAAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTGGC 2042
Db
2006 ACCTGCTATTCGGTGTGTGTGAAACACAGTGTCTCGTGTACCTCTCGTGGAGTGGC 2065
QY
2043 AACTGATGAACAGAGAAAGTTAAATACAGATGTTTTTCCAAAGAACTCTTGACCA 2102
Db
2066 AACTGAAGAACAGAGAAAGTTAAATACAGATGTTTTTCTAAAGAACTCTTGACCA 2125
QY
2103 TGACAGATGTGACAGCACACAGATTTTACAGCTGTACAGCCACACCAATGACTGCA 2162
Db
2126 TGACAGATGTGACAGCACACAGATTTTACAGCTGTACAGCCACACCAATGACTGCA 2185
QY
2163 CTGCTCAATGACCAATTTGTTCCCGAGAACACACAGCTCTCAGAGGCGCAGATCTCCAT 2222
Db
2186 CTGCTCAATGATCACTGTGCTGCTGAAACACACAGCTGTGACAGAGGCGCAGATCTCCAT 2245
QY
2223 TTTTAGGTATGAAATTTGCTGCAAGTAAACCCCAAGTAACTGTACTGTAAACAGAGACAG 2282
Db
2246 TGCCAAAGTATGAGATTTGCTGCAAGTAAACCCCAAGTAACTGTACTGTAAACAGAGACAG 2305
QY
2283 CTGACAGAGCTGTGCTGCTGGAACACAGAACTGCGCAGTGGGAGCCCGGAATCAGAGTGCAT 2342
Db
2306 CTGACAGAGCTGTGCTGCTGGAACACAGAACTGCGCAGTGGGAGCCCGGAATCAGAGTGCAT 2365
QY
2343 TGCCCTGCGCGAAATATCTGTGGCATTTGCTGGCATTTGGTGGAACTCAATGTTGAA 2402
Db
2366 CGCCCTGCGCGAAATATCTGTGGCAATTTGCTGGCAATTTGGTGGAACTCAATGTTGAA 2425
QY
2403 AATTACTACTGCGCAAGGAAATTTGCAATGCTAAATGTTCTGTAGGAAACCAATGTC 2462
Db
2426 AATCACTACTGCTAAGGAAATTTGCAATGCTAAATGTTCTGTAGGAAACCAATGTC 2485
QY
2463 CTTTGTGGCTTCTTCTTAACCCAGAGAGATGTTGCTTCCCTTAAGCAGCTTCGAT 2522
Db
2486 CTTTGTGGCTTCTTCTTCCATCCAGAGAGGTTGGAGTGTGCTTAAAGCAGCTTCGAT 2545
QY
2523 AATGAGTCACTCTCAGAGCATGTCAAGCTCACTTAAACCCAGTGGTGGCTTCGAA 2582
Db
2546 AATGCAATCTCTCAAGTATGTTCCAGCTCACTCTGACTCCATGGTGGTCTTCGAA 2605
QY
2583 GATCAATGTCTCTACTGTGTGGAGAGATATGTTCCCATTTACAAATAGTTTACTACA 2642
Db
2606 GATCAATGTCTCTACTGTGTGGAGAGATATGTTCCCATTTACAAATAGTTTACTACA 2665
QY
2643 GTGATGCGCTCTGAGCCAGTGTGATGCTGGAATTTTATCAGAACCCAGTAC 2702

Db 673 GAGAACTAACTGGCTCTCTCGATTGTAAACAGATGGACCTGGGAATTTAATAATAAGA 732
QY 467 CGAGTGCAGCTGGCTCATTTGAGGACACGCCAAATAGAAATAGACATTTGTTTCATC 526
Db 733 CGAGTGCACATGGCTCATTTGAGGACACGCCAAATAGAAATAGACATTTGTTTCATC 792
QY 527 ATTITGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATG 586
Db 793 ATTITGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATG 852
QY 587 CACCGTAGTGTCTGATTTAGTGGCTCATTTGTTCCCTGAGAGAGATGGCAATGAGACTG 646
Db 853 CACCTCTGATTGCTGCTTCTAGTGGCTCATTTGTTCCCTGAAAGAGATGGCAATGAGACGG 912
QY 647 TCCCTGAGGTTGTTGCCACATCAGCTTATGCTTCTGCTGCATTTTTTTAGTGCATGCTGTT 706
Db 913 CTCTGAGGCTACTGCTACTTCAGGTTATGCACTGCTGCAATTTTTTCAATGATGCTGCTT 972
QY 707 ATAAITGAGTGGATTAATATTAATTAAGTTTATGATGTTTATGATGTTTATGATGTTTATG 766
Db 973 ATAAITGAGTGGATTAATATTAATTAAGTTTATGATGTTTATGATGTTTATGATGTTTATG 1032
QY 767 GCCGAGGAGAGTGAAGATCAGTAAATAGCAGGAACTGTTGAATGTAATGTTCTGAA 826
Db 1033 GCCGAGGAGAGTGAAGATCAGTAAATAGCAGGAGCTGTTGAGTGAATGTTCTGAA 1092
QY 827 ACTGAAAGGTGAAGCATGTGCAATTCCTCACTGATACAGAACTGTGTTTCTCTCATC 886
Db 1093 ACTGAAAGGTGAAGCATGTGCAATTCCTCACTGATACAGAACTGTGTTTCTCTCATC 1152
QY 887 GAGGATCTGCATTTCAAGTGAATGTCAGAGATGCTCTCTGTTCTCAGACTGGCAGGTC 946
Db 1153 GAGGATCTGCATTTCAAGTGAATGTCAGAGATGCTCTCTGTTCTCAGACTGGCAGGTC 1212
QY 947 CTGGATGTTCACTTCTCTGACAGTAACTGATGTTGACCTGAGAGAAATATTCTA 1006
Db 1213 CTGGATGTTCACTTCTCTGACAGTAACTGATGTTGACCTGAGAGAAATATTCTG 1272
QY 1007 ACTTAAAGCTCCAGAGATCTCAATGAGTGTGTTCAATGAGAAATATTGTTGGTTG 1066
Db 1273 ATTTAAAGCTCCAGAGATCTCAATGAGTGTGTTCAATGAGAAATATTGTTGGTTG 1332
QY 1067 TTGAGGATATATGTTCAACACTCAGATTAATACATGTTTCTAGCGTATGACCTGCTT 1126
Db 1333 TTGCGGATATATGTTCAACACTCAGATTAATACATGTTTCTAGCGTATGACCTGCTT 1392
QY 1127 CTAGGAGTGTCTTCCACTAAACCGTTCTGTGAACATGTTGTTGTTAGATATGTTCAAT 1186
Db 1393 CTAGGAGTGTCTTCCACTAAACCGTTCTGTGAACATGTTGTTGTTAGATATGTTCAAT 1452
QY 1187 CTTTGGCATTAACAAGATAAATTTACATGTTATGAGGAAATTTGATCCACTGGA 1246
Db 1453 CTTTGGCATTAACAAGATAAATTTACATGTTATGAGGAAATTTGATCCACAGGA 1512
QY 1247 ATGTGACCAATGAGTGTGAGAGTTTTCATTTCAATTAATGAGTCAATGGTGTGTTGACCC 1306
Db 1513 ACGTGACCAATGAGTGTGAGAGTTTTCATTTCAATTAATGAGTCAATGGTGTGTTGACCC 1572
QY 1307 CTAGGCAAGAGGAGTATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1366
Db 1573 CGAAGCTTAAGGATCAGTATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1632
QY 1367 ATGCGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426
Db 1633 CTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1692
QY 1427 ATGTGCGAGGAATATGATTGGATTAAGAACACATGAGTATATTACACACCGAGGTGCC 1486
Db 1693 TTGTGCGAGGAATATGATTGGATTAAGAACACATGAGTATATTACACACCGAGGTGCC 1752
QY 1487 TTGTGCGAGGAGTATGAGGATGAGTGTGTTTACGACCATAGGACCGAGGCTTATAG 1546
Db 1753 TTGTGCGAGGAGTATGAGGATGAGTGTGTTTATGATGAGGACCGAGGCTTATAG 1812

QY 1547 TTATGCTGCTACAGGCTTTTCAGTGGCAATTAAGTACCGGCTTGAGAGATGATCTCTACC 1606
Db 1813 TTATGCTGCTACAGGCTTTTCAGTGGCAATTAAGTACCGGCTTGAGAGATGATCTCTACA 1872
QY 1607 GATATGATGAGTATACCCAGATGTGACCAATTTCTTAAGGACAGCCGATTTTCCGTTACT 1666
Db 1873 GATACGATGAGTATACCCAGATGTGACCAATTTCTTAAGGACAGCCGATTTTCCGTTACT 1932
QY 1667 TGACACAGCTGTGATGATGAGTGGAAACCAATGCTGTTGTTGGGGGAAACACACACAATG 1726
Db 1933 TGACACAGCTGTGATGATGAGTGGAAACCAATGCTGTTGTTGGGGGAAACACACACAATG 1992
QY 1727 ACACATCTATGAGCAGTGGCCAAATGCTTCTTTACAGATTTTCAATGCTTATGATGATG 1786
Db 1993 ACATTTCCATGAGCAGCTGGTCCAAATGCTTCTCTGAGACTTCAATGCTTATGATGATG 2052
QY 1787 CTGTGACCGCTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACACAGATTTG 1846
Db 2053 CTGTGACCGCTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACACAGATTTG 2112
QY 1847 GCCATTCAGAGTCTTACACACAGCAGCAGTATGTTGTTGCTGTTTCAATGATCTCTC 1906
Db 2113 GCCATTCAGAGTCTTACACACAGCAGCAGTATGTTGTTGCTGTTTCAATGATCTCTC 2172
QY 1907 TCCTCAGCAGCATCTCTGTTATTTCACTCGAAACAGTGTGATGCGCATCGAGTGAAGCG 1966
Db 2173 TCCTCAGTCACTCTGTTTACCTCGAGCAGTGGATGCAACACCGCAGTGAAGCTG 2232
QY 1967 CTGTTTACAGCAGGACCTGTTATTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2026
Db 2233 CTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2292
QY 2027 TCTCTGCGGCTGGCACTGATGAACAGAAAGTTAAATCAGAAATCTGTTTCTTCA 2086
Db 2293 CTCTCTGGAGTTGGCACTGAGAAACAGCAAGAAAGTTAAATCAGAGTGTGTTTCTA 2352
QY 2087 AAGAACTCTTGACCATGACATGTGACCAAGCAACAGATTTTACAGTGTGTAAGCA 2146
Db 2353 AAGAACTCTTGACCATGACATGTGACCAAGCAACAGATTTTACAGTGTGTAAGCA 2412
QY 2147 ACACCAATGACTGCCACTGCTCAATGACCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2206
Db 2413 ATACCAATGACTGCCACTGCTCAATGATCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2472
QY 2207 AAGCCGAGATCTCCATTTTATGATGAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2266
Db 2473 AAGCCGAGATCTCCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2532
QY 2267 GTACCAAGAGACAGCTGCTGAGGCTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2326
Db 2533 GCAATTAAGAAACAGCTGCTGAGGCTGTTGCTTACAGCAACTGCTGTTGTTGTTGTTGTTG 2592
QY 2327 GGAATCAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2386
Db 2593 GGAATCAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652
QY 2387 GAACTCATGTTTGAATAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2446
Db 2653 GAACTCATGTTTGAATAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2712
QY 2447 GTAGGAAACCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2506
Db 2713 GTAGGAAACCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2772
QY 2507 TTAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2566
Db 2773 TTAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2832
QY 2567 GGGTGGCTTGGGAGATCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2626
Db 2833 GGGTGGCTTGGGAGATCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2892

```
QY 2627 CAAATAGTTTACTACAGTGGATCGGCTGAGCCAGTGATGCTGGATTCCTGTGGAATTT 2686
DB 2893 CAAATAGTTTCTGAGTGGATGCAATCTGAGCCAGTGATGCTGGCTTCTGTGGATCT 2952
QY 2687 TATCAGAACCCAGTACTCGGGAGTGAAGGCTGCAACCTGTCATCAACCCTCAATGGTA 2746
DB 2953 TGTGAGAGCTAGTACTCGGGGATTAAGGCTGCAACCTGTCATCAACCCTCAATGGCA 3012
QY 2747 GTGCTGTGAAAGGCTGCAACACACAGTGTAGCAGTCCGAGACCACTGTGCCCTGA 2806
DB 3013 GCGTCTGTGAAAGGCTGCAACACACAGTGTAGCAGTCCGAGACCACTGTGCCCTGC 3072
QY 2807 GGACAGCATGTGGAGATTGCACACAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGA 2866
DB 3073 GGACAGCATGTGGAGTGCACAGCAGCTCGGAGTGCATGTGGTGCAGTAAATGA 3132
QY 2867 AGCAGTGTGGACATCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGAAT 2926
DB 3133 AGCAGTGTGGACATCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGAAT 3192
QY 2927 GGTATACGATGAGCACCTGCCCTGAAATTTGTTAGGCTACTGTACCTGTAGTCATT 2986
DB 3193 GGTATACGATGAGCAGTCCCACTGAAATTTGTTAGGCTACTGTACCTGTAGTCATT 3252
QY 2987 GCTTGGACCAACAGGCTGTGGCTGTACTGATCCAGCAATCTGCGCAAGGAAAT 3046
DB 3253 GCTTGGACGAGCAGGCTGTGGCTGTACTGATCCCTAGCAATCTGCGCAAGGAAAT 3312
QY 3047 GGTATAGAGGTTCTTAAAGGACACAGTGAAGTCCCTTGCAGCCCTCACAGAAAT 3106
DB 3313 GTATTGAGGCGCAGTATAAGGACCTGTGAAGATCCCTGCAGCCCTCTCGAGAAATG 3372
QY 3107 TCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGACACAGATACAACTGGT 3166
DB 3373 TGTATCCACAGCCCTCTGTAACCTCCAGCATGTGTCTAGAGACACAGATACAACTGGT 3432
QY 3167 CTTTCAATTCAGTCCAGCTTCCAAATGCAACGCGCACAGTAAATGCATCAATCAGACA 3226
DB 3433 CTTTCAATTCAGTCCAGCTTCCCAAGTCAACGCGCACAGCAATGCATCAACAGAGTA 3492
QY 3227 TCTGTGAGAGTGTGAGACCTTGACACAGGCAAGCACTGCAGACCTGCATATCTGGCT 3286
DB 3493 TCTGTGAGAGTGTGAGACCTTGACACAGGCAAGCACTGCAGACCTGCATATCTGGCT 3552
QY 3287 TCTAGCGTATCCCAATGAGGGAATGTGAGCCATGCAAGTGCATGAGGCAAGCGT 3346
DB 3553 TCTATGTTGACCCGACTAATGAGGCAAGTGTGAGCCATGCAAGTGCATGAGGCAAGCGT 3612
QY 3347 CTCTGTGCAACCAACACAGGCAAGTGTCTTGCACCAACAGCGCGTCAAGGGGAGC 3406
DB 3613 CACTGTGCAACCAACACAGGCAAGTGTCTTGCACCAACAGCGCGTCAAGGGGAGC 3672
QY 3407 AGTGCAGCTATGTGAGGTAGAAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTT 3466
DB 3673 AGTGCAGCTATGTGAGGTAGAAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTT 3732
QY 3467 ATTATACCTCTCTTATGATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCT 3526
DB 3733 ACTATACCTCTCTTATGATCAGTTCACCTTTAGTCTATCCAGGAGGAGCAGCGCT 3792
QY 3527 ATTACACAGCTATCAATTTTGTGGCTACTCTGACGAACAAACAGGAGATTGGAATGT 3586
DB 3793 ACTACACAGCTATCAATTTTGTGGCTACTCTGATGAAACAAACAGGAGATTGGAATGT 3852
QY 3587 TCATCAATGCTCTCAAGAAATTTCAACCTCAACATCAACCTGGCGTGCAGTTTCTCAGCTG 3646
DB 3853 TCATCAATGCTCTCAAGAAATTTCAACCTCAACATCAACCTGGCGTGCAGTTTCTCAGCTG 3912
QY 3647 GAACCCAGGCTGAGAGAGATGCTGTTTTCACCAACCAACCAATTAAGGAGTACAAG 3706
DB 3913 GAACCCAGATGAGAGAGAGTGCCTGTTTTCACCAACCAACCAATTAAGGAGTACAAG 3972
QY 3707 ATAGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACCAACCAATATCACTTTCTTTGTTT 3766
```

```
DB 3973 ATAGTCTCTCTAATGAGAATTTGAATTTTCGCAACCAACCAATCACTTTCTTTGTTT 4032
QY 3767 ATGTCAGTAAATTCACCTGCGCCATCAAAATTCAGGT 3803
DB 4033 ATGTCAGTAAATTCACCTGCGCCATCAAAATTCAGAT 4069

RESULT 12
AB093245 5693 bp mRNA linear ROD 07-FEB-2003
LOCUS Mus musculus mRNA for mKIAA0548 protein.
DEFINITION AB093245
ACCESSION AB093245
VERSION AB093245.1 GI:26006172
KEYWORDS FLI CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T.,
Ohara,O. and Koga,H.
Prediction of the coding sequences of mouse homologues of KIAA
gene: I. The complete nucleotide sequences of 100 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
Unpublished
2 (bases 1 to 5683)
Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
Direct Submission
Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamata-ri, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
Location/Qualifiers
1..5683
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg03041"
/tissue_type="brain"
/dev_stage="adult"
/notes="vector:modified pBC SK-"
1..5683
/gene="mKIAA0548"
<1..3181
/gene="mKIAA0548"
notes="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=2
/evidence=not experimental
/product="mKIAA0548 protein"
/db_xref="GI:26006173"
/translation="AYDLTSREWLPLNHSVSVVRYGSHLSLHDKIYMYGKIDST
GNVNEURPHIHNSWLLTPKAKQYAVGHSALHTLASGRVYMLVIFGCPLYG
YISVYQBYDLEKNTWSILTGALVGGYGHSSVYDRTKALIVHGG/KAFSANKVRL
ADLIYDYDITQMTTLKDSRFFRYLHTAVIVSGTVMVFGNTNLLSMGHCAPSS
DFMAYIDCDRWSVLPRPELHDVNRFGHSALVYNSMTYVFGFNSLLSDVLFVLE
QCDARSEACAAAGPGIRCLDQTSRCTLSEWELATBEQAEKLSCEFKRTLDHRC
DQHTDCYCTANTNDCHWCDHCVPNHSCTEGQISIAKYESCPKDNPMYCNKTKSC
RSCALDQNCWEPNRQECIALPENI CGNGHLVGNCLKITAKENYDNALSCRNH
AFIALDTSQKKVFLKQLMQSSQSKSLTLPWGLKINVSYWCHWEDMSPFTNS
LLQMPSPEDAGFCGILSEFTRGLKAATCINPLGSCVETTPANHSKQRCPTCALR
TACCECTSSSECMWCSNMKQCVDSNAYVASFPQCMWEYTSVSCPCCTGYCTCS
HCLPQCGCTDPTDPTGKGTGEGYKGPVAMPQASAGNVPQPLNLSMCLDR
YNSFIHCPACQCNHSGKSNQSVCEKEDLTTCGHCTCTISGFGPTNGKQCPCK
CNGHSLCNTNTGKCFCTTKGVGDECOLCEVENRYGNPLKGTTCVYTLIDYQTFPS
LSQEDDRYTAINFVTPDEQNRLDNFINASKNFNLNITWATSPFAGTGTGEVPV
```


SKTIIKEYKDSFSENEKDFRNHPNITFFVVSNTFWPIKIQIAFQSHNSFMDLVQFFV
TFPSCSLILLVAUVVKIKQCSWASRRREQLLREMOQWASPPASVNVVALETDREPP
DLIGSIIKTVKPELALPEPCFNGKAAVLVSFVRLPRGLGGIPPGSGGLAVASALVDIS
QQMPIVYKESGAVNRKQPPAQPGTCTI"

ORIGIN

Query Match 59.6%; Score 2276.4; DB 10; Length 5683;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 2433; Conservative 0; Mismatches 261; Indels 0; Gaps 0;

QY	1110	ACGGTATGACCTTCTTACGGAGTGGCTTCCACTAAACCGTTCCTGCAACAATGTGGT	1169
DB	1	ACGGTATGACCTTCTTACGGAGTGGCTTCCACTAAACCGTTCCTGCAACAATGTGGT	60
QY	1170	TGTTAGATATGGTCATTCTTTGGCATTATACAGGATMAAATTTACATGTATGGAGGAAA	1229
DB	61	TGTAAGATATGGTCATTCTTTGGCATTATACAGGATMAAATTTACATGTATGGAGGAAA	120
QY	1230	AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTCATTAATGATC	1289
DB	121	AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTCATTAATGATC	180
QY	1290	ATGGGTGTTGTTGACCCCTTAAGCCAAAGAGGAGGATGAGTGGTGGGCACTCTGCACA	1349
DB	181	ATGGGTGTTGTTGACCCCTTAAGCCAAAGAGGATGAGTGGTGGGCACTCTGCACA	240
QY	1350	CATTGTTACATGGAAGAAATGGCGAGTGGTCAATGCTGTTTGGTCACTGCCCTCT	1409
DB	241	CATTGTTACATGGAAGAAATGGCGAGTGGTCAATGCTGTTTGGTCACTGCCCTCT	300
QY	1410	CTATGGATATAAGCAATGTCAGGAATATGATTTGGATAGACACATGAGTATATT	1469
DB	301	CTATGGATATAAGCAATGTCAGGAATATGATTTGGATAGACACATGAGTATATT	360
QY	1470	ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACCGGCATAGCAGTGTTCACGACCATAG	1529
DB	361	ACATACCTCAGGGTGCTCTTGTGCAAGGGGGTTATGCCACAGTAGTGTTCATGATGACAG	420
QY	1530	GACGAGGCGCTATAGCTTCATGTTGCTACAGGCTTTAGTGGCCATTAAGTACCGGCT	1589
DB	421	GACGAGGCGCTATAGCTTCATGTTGCTACAGGCTTTAGTGGCCATTAAGTACCGGCT	480
QY	1590	TGCAGATGATCTCTACCGATATGATGTGATACCGATGAGGACCATTTCTTAAGGACAG	1649
DB	481	TGCAGATGATCTCTACCGATATGATGTGATACCGATGAGGACCATTTCTTAAGGACAG	540
QY	1650	CGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGAAACCAATGCTGTTGTTGG	1709
DB	541	CGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGAAACCAATGCTGTTGTTGG	600
QY	1710	GGGAACACACACATGATCATCTATGAGCCATGCGCCCAATGCTTCCTTCAGATTT	1769
DB	601	GGGAACACACACATGATCATCTATGAGCCATGCGCCCAATGCTTCCTTCAGATTT	660
QY	1770	CATGCCCTATGACATGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA	1829
DB	661	CATGCCCTATGACATGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA	720
QY	1830	TGATGTCAACAGATTTGGCCATTAGCAGTCTTACACAGCACCAGTATGTTGCG	1889
DB	721	TGATGTCAACAGATTTGGCCATTAGCAGTCTTACACAGCACCAGTATGTTGTTGCG	780
QY	1890	TGGTTTCAATAGTCTCTCTCTCAGCAGCATCTGATTCACCTCGGAACAGTGTATGC	1949
DB	781	TGGTTTCAATAGTCTCTCTCTCAGCAGCATCTGATTCACCTCGGAACAGTGTATGC	840
QY	1950	GCATCGGAGTGAAGCGCTTGTGATGACGAGGACCTGTTATTCGGTGTGTGTGGAAACAC	2009
DB	841	GCATCGGAGTGAAGCGCTTGTGATGACGAGGACCTGTTATTCGGTGTGTGTGGAAACAC	900
QY	2010	AGGGTCTGCTCAGTGTATCTCTGTTGGGCGCTGGCACTGATCAACAGAAAGATTAAA	2069
DB	901	AGGGTCTGCTCAGTGTATCTCTGTTGGGCGCTGGCACTGATCAACAGAAAGATTAAA	960

QY	2070	ATCAGATGTTTTCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATTG	2129
DB	961	ATCAGATGTTTTCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATTG	1020
QY	2130	TTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTGCATATGATGTGCCCCAG	2189
DB	1021	TTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTGCATATGATGTGCCCCAG	1080
QY	2190	GAACACAGCTGTGACAGGCGCAGATCTCCATTTTGTAGTATGAGAAATGCCCCAAGGA	2249
DB	1081	GAACACAGCTGTGACAGGCGCAGATCTCCATTTTGTAGTATGAGAAATGCCCCAAGGA	1140
QY	2250	TACCCCATCTACTACTGTAAACAAGAACCACTGCGAGAGCTGCGCCCTGACACAGAA	2309
DB	1141	TACCCCATCTACTACTGTAAACAAGAACCACTGCGAGAGCTGCGCCCTGACACAGAA	1200
QY	2310	CTGCCAGTGGAGCCCGGGAATCAGGAGTGCAATGCGCTGCCGCGGAAATATCTGGGCAT	2369
DB	1201	CTGCCAGTGGAGCCCGGGAATCAGGAGTGCAATGCGCTGCCGCGGAAATATCTGGGCAT	1260
QY	2370	TGGCTGCACTTGTGGAACTCATGTTTGAATAATTAATCTACGCAAGAGAAATATGA	2429
DB	1261	TGGCTGCACTTGTGGAACTCATGTTTGAATAATTAATCTACGCAAGAGAAATATGA	1320
QY	2430	CAATGCTAAATGTTCTGTAGGAACCAATGCGCTTTTGGCTTCTTTACACCCAGAA	2489
DB	1321	CAATGCTAAATGTTCTGTAGGAACCAATGCGCTTTTGGCTTCTTTACACCCAGAA	1380
QY	2490	GAGGTAGATTTGCTTAAAGCAGCTGCAATTAATGAGTCACTCTCAGAGCATGTCCAA	2549
DB	1381	GAGGTAGATTTGCTTAAAGCAGCTGCAATTAATGAGTCACTCTCAGAGCATGTCCAA	1440
QY	2550	GCTCAGCTTAAACCCCATGGCTCGGCTTCGGAAGATCAATGTCTCTACTGTGTGGGA	2609
DB	1441	GCTCAGCTTAAACCCCATGGCTCGGCTTCGGAAGATCAATGTCTCTACTGTGTGGGA	1500
QY	2610	AGATATGTCCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2669
DB	1501	AGATATGTCCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1560
QY	2670	TGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAAGGCTGCAACCTGCAT	2729
DB	1561	TGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAAGGCTGCAACCTGCAT	1620
QY	2730	CAACCCCTCAATGGCAGCGTCTGTGAAAGGCTGCAACCCAGTGTCTAAGCAGTGGCG	2789
DB	1621	CAACCCCTCAATGGCAGCGTCTGTGAAAGGCTGCAACCCAGTGTCTAAGCAGTGGCG	1680
QY	2790	GACACCATGTGCTTGGAGCAGCATGTGGAGATTGCAACCGGCGCAGCTCTGAGTGCAT	2849
DB	1681	GACACCATGTGCTTGGAGCAGCATGTGGAGATTGCAACCGGCGCAGCTCTGAGTGCAT	1740
QY	2850	GTGGTGCAGCAACATGAAGCAGTGTGACCTTCAATGCTATGTGGCTTCTTCCCTTT	2909
DB	1741	GTGGTGCAGCAACATGAAGCAGTGTGACCTTCAATGCTATGTGGCTTCTTCCCTTT	1800
QY	2910	TGGCCAGTGTATGGAAATGGTATACGATGAGCAGCTTCCGCGGCTGAAATTTGTAGGCTA	2969
DB	1801	TGGCCAGTGTATGGAAATGGTATACGATGAGCAGCTTCCGCGGCTGAAATTTGTAGGCTA	1860
QY	2970	CTGTACTCTGTAGTCAATGCTTGGAGCAACAGCGCTGTGGCTGGTGTACTGTATCCACAGAA	3029
DB	1861	CTGTACTCTGTAGTCAATGCTTGGAGCAACAGCGCTGTGGCTGGTGTACTGTATCCACAGAA	1920
QY	3030	TACTGTCAAGGAAATGGATAGAGGTTCTTAAAGGAAACAGTGAAGATGCTTCCGCA	3089
DB	1921	TACTGTCAAGGAAATGGATAGAGGTTCTTAAAGGAAACAGTGAAGATGCTTCCGCA	1980
QY	3090	AGCCCTCTACAGGAAATTTCTATCCACAGCCTCTCAATTCACAGCATGTGTCTAGAGGA	3149
DB	1981	AGCCCTCTACAGGAAATTTCTATCCACAGCCTCTCAATTCACAGCATGTGTCTAGAGGA	2040


```
QY 3150 CAGCAGATACAACTGGTCTTTTCAATTCAGTCCAGCTTGCATGCAATGCAACCGGCCACAGTAA 3209
Db 2041 CAGCAGATACAACTGGTCTTTTCAATTCAGTCCAGCTTGCATGCAATGCAACCGGCCACAGTAA 2100
QY 3210 ATGATCATCAATCAGACGATCTGTGAGAGTGTGAGAACTGACACAGCGCAAGCACTGCGA 3269
Db 2101 ATGATCATCAACAGAGTGTGTGAGAGTGTGAGAACTGACACAGCGCAAGCACTGCGA 2160
QY 3270 GACCTGCATATCTCGCTTTCTACGGTGTATCCACCAATGAGGAGGAAATGTGAGCAATGCA 3329
Db 2161 GACCTGCATATCTCGCTTTCTATGTGACCCGACTAATGGAGGCAATGTGAGCAATGCA 2220
QY 3330 GTGCAATGGGCAACGGTCTCTGTGGAACACCAACAGCGGCAAGTGTCTTGTGACCAACAA 3389
Db 2221 GTGCAATGGGCAACGGTCTCTGTGGAACACCAACAGCGGCAAGTGTCTTGTGACCAACAA 2280
QY 3390 GGGCGTCAAGGGGACAGTCCAGCTATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3449
Db 2281 AGGTGTCAAGGGGACAGTCCAGCTATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2340
QY 3450 TCTCAGAGGAACATGTTTATTAATATATCTCTTCTTATTAATATATCAGTGTGAGTGTGAGT 3509
Db 2341 TCTCAGAGGAACATGTTTATTAATATATCTCTTCTTATTAATATATCAGTGTGAGTGTGAGT 2400
QY 3510 CCAGGAACATGTCCTTATTAACAGCTATCAATTTTGTGCTACTCTCTGACGAAACAAA 3569
Db 2401 CCAGGAACATGTCCTTATTAACAGCTATCAATTTTGTGCTACTCTCTGACGAAACAAA 2460
QY 3570 CAGGATTTGACATGTTTATTAATATATCTCTTCTTATTAATATATCAGTGTGAGTGTGAGT 3629
Db 2461 CAGGATTTGACATGTTTATTAATATATCTCTTCTTATTAATATATCAGTGTGAGTGTGAGT 2520
QY 3630 TGCCAGTTTCTCAGCTGAAACCCAGCTGAGGAGAGAGTGTGTTTTCAGAAACCAA 3689
Db 2521 CACCAGCTTCCAGCGGAAACCCAGCTGAGGAGAGAGTGTGTTTTCAGAAACCAA 2580
QY 3690 CATTAAGAGATACAAAGATAGTTTCTCTAATGAGAGTGTGATTTTCGCAACCCCAAAA 3749
Db 2581 CATTAAGAGATACAAAGATAGTTTCTCTAATGAGAGTGTGATTTTCGCAACCCCAAAA 2640
QY 3750 TATCACTTTCTTTTGTATGACGATTAATTCACCTGGCCCAATCAAAATTCAGGT 3803
Db 2641 CATCACTTTCTTTTGTATGACGATTAATTCACCTGGCCCAATCAAAATTCAGAT 2694

RESULT 13
AR164818
LOCUS AR164818 2625 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 18 from patent US 6274339.
ACCESSION AR164818
VERSION AR164818.1 GI:16238029
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2625)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 18 14-AUG-2001;
FEATURES Location/Qualifiers
source 1..2625
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 58.9%; Score 2248.2; DB 6; Length 2625;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

QY 3 GGTGGCCGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGGCGGCGGCGGCGGCGGCGG 62
Db 9 GGTGGCCGCGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGGCGGCGGCGGCGGCGGCGG 68
```

```
QY 63 GTTCGGGCGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122
Db 69 GTTCGGGCGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128
QY 123 GGGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
Db 129 GGAACCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 188
QY 183 GTGTGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 242
Db 189 GTGTGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 239
QY 243 CGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302
Db 240 CGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 299
QY 303 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 362
Db 300 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 359
QY 363 CGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 422
Db 360 CGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 419
QY 423 TTCTGGGTTGTGACAGATGGACCTGCAAAATATATAATACAAAACGAAAGTCACGCTGGCT 482
Db 420 TTCTGGGTTGTGACAGATGGACCTGCAAAATATATAATACAAAACGAAAGTCACGCTGGCT 479
QY 483 CATTAAGAGCAGCCAAATAGATAATAGACTTCTGTTTCAATCATTTTGTCTACAGAGTG 542
Db 480 CATTAAGAGCAGCCAAATAGATAATAGACTTCTGTTTCAATCATTTTGTCTACAGAGTG 539
QY 543 TAGTTCGGGACCAATTAATATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGC 602
Db 540 TAGTTCGGGACCAATTAATATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGC 599
QY 603 ATTTAGTGGCTCATTTGTTCTTGAGAGAGATGSCAATGAGACTGTCCCTGAGGTTGTTC 662
Db 600 ATTTAGTGGCTCATTTGTTCTTGAGAGAGATGSCAATGAGACTGTCCCTGAGGTTGTTC 659
QY 663 CACATCAGGTTATGCTTGTCTGCTGCTATTTTGTAGTGTGCTTATTAATTTGACTGGATT 722
Db 660 CACATCAGGTTATGCTTGTCTGCTGCTATTTTGTAGTGTGCTTATTAATTTGACTGGATT 719
QY 723 TAATATTACTTACAGTTTGTGATATGTCCAAATAACTGCTCAGGCGGAGGAGAGTGTAA 782
Db 720 TAATATTACTTACAGTTTGTGATATGTCCAAATAACTGCTCAGGCGGAGGAGAGTGTAA 779
QY 783 GATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTGTTGAAAACCTGGAAGAGTGAAGC 842
Db 780 GATCAGTAATAGCAGCGAATACTGTTGAATGTGAATGTGTTGAAAACCTGGAAGAGTGAAGC 839
QY 843 ATGTGACATTCCTCAGTGTACAGCAACTGCTGCTTTTCTCATCGAGGCAATCTGCAATTC 902
Db 840 ATGTGACATTCCTCAGTGTACAGCAACTGCTGCTTTTCTCATCGAGGCAATCTGCAATTC 899
QY 903 AAGTGTGATGTGAGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCTGAGATGTTGAGTTCC 962
Db 900 AAGTGTGATGTGAGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCTGAGATGTTGAGTTCC 959
QY 963 TGTACAGCTTAACAGCTCATTTTGGACTCGAGAGGATATTTCTAACTTAAAGCTCCCCAG 1022
Db 960 TGTACAGCTTAACAGCTCATTTTGGACTCGAGAGGATATTTCTAACTTAAAGCTCCCCAG 1019
QY 1023 AGCATCTCATAAAGCTGTGTGCTCAATGGAAACATTATGTGCGGTTGTTGGAGGATATATGTT 1082
Db 1020 AGCATCTCATAAAGCTGTGTGCTCAATGGAAACATTATGTGCGGTTGTTGGAGGATATATGTT 1079
QY 1083 CAACCACTCAGATTAATACATGTTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCC 1142
Db 1080 CAACCACTCAGATTAATACATGTTTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCC 1139
```

1143 ACTAAACCGTTCTGTGAACAATGTGGTGTGTAGATATGTCATCTTTGGCATTTATACAA 1202
1140 ACTAAACCGTTCTGTGAACAATGTGGTGTGTAGATATGTCATCTTTGGCATTTATACAA 1199
1203 GGATAAAATTTACATGTATGGAGGAAAATTCATCCAACTGGGAATGTGACCAATGAGTT 1262
1200 GGATAAAATTTACATGTATGGAGGAAAATTCATCCAACTGGGAATGTGACCAATGAGTT 1259
1263 GAGAGTTTTCATCATTAATGAGTCATGGGTGTGTGATGAGTGTGATGAGTGTGATGAGT 1322
1260 GAGAGTTTTCATCATTAATGAGTCATGGGTGTGTGATGAGTGTGATGAGTGTGATGAGT 1319
1323 GTATGAGAGTGTGGGACCTCTGCACATTTCTACATGAGATGCGGAGTGGTCAT 1382
1320 GTATGAGAGTGTGGGACCTCTGCACATTTCTACATGAGATGCGGAGTGGTCAT 1379
1383 GCTGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGAGGAATATGA 1442
1380 GCTGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGAGGAATATGA 1439
1443 TTTGGATAGACACATGAGTATATACACACCCAGGTCCTTGTGCAAGGGGTTA 1502
1440 TTTGGATAGACACATGAGTATATACACACCCAGGTCCTTGTGCAAGGGGTTA 1499
1503 CGGCCATAGCAGTGTACACCATAGGACACAGGCGCTTATACGTTCTATGTGGCTACAA 1562
1500 CGGCCATAGCAGTGTACACCATAGGACACAGGCGCTTATACGTTCTATGTGGCTACAA 1559
1563 GGCCTTCAGTGCAATAGTACCGGCTTGCAGATGATCTTACCGATATGATGGATAC 1622
1560 GGCCTTCAGTGCAATAGTACCGGCTTGCAGATGATCTTACCGATATGATGGATAC 1619
1623 CCAGATGTGACCATCTTTAAGGACAGCGATTTTTCGTTTCTTGTGACACAGCTGTGAT 1682
1620 CCAGATGTGACCATCTTTAAGGACAGCGATTTTTCGTTTCTTGTGACACAGCTGTGAT 1679
1683 AGTAGTGAAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1742
1680 AGTAGTGAAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1739
1743 TGGCCCAAAATCTTCTTCAGATTTTCAGGCTATGACATGTCGCTGTGACCCCTGGTC 1802
1740 TGGCCCAAAATCTTCTTCAGATTTTCAGGCTATGACATGTCGCTGTGACCCCTGGTC 1799
1803 AGTGTCTCCGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862
1800 AGTGTCTCCGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1859
1863 ACACAGACCATGATGTGTTGGTGGTGTGATGATGATGATGATGATGATGATGATGAT 1922
1860 ACACAGACCATGATGTGTTGGTGGTGTGATGATGATGATGATGATGATGATGATGAT 1919
1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTTTGTAGCAGCAGG 1982
1920 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTTTGTAGCAGCAGG 1979
1983 ACCTGGTATTCGGTGTGTGGAACACAGGTCGTCTCAGTGATCTCCTGCGGCGCTGGC 2042
1980 ACCTGGTATTCGGTGTGTGGAACACAGGTCGTCTCAGTGATCTCCTGCGGCGCTGGC 2039
2043 AACTGATCAACAGAGGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCA 2102
2040 AACTGATCAACAGAGGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCA 2099
2103 TGACAGATGTGACCAAGCAACAGATTTGTTACAGCTGTACAGCCCAACCAATGATGCCA 2162
2100 TGACAGATGTGACCAAGCAACAGATTTGTTACAGCTGTACAGCCCAACCAATGATGCCA 2159
2163 CTGGTGCAATGACCATGTGTCCTCCAGGACACAGCTGCTCAGAGCCAGATCTCCAT 2222
2160 CTGGTGCAATGACCATGTGTCCTCCAGGACACAGCTGCTCAGAGCCAGATCTCCAT 2219
2223 TTTTAGTATGAGAATTTGCCCCCAAGGATAACCCCATGTACTGTGTAAACAAGAACCCAG 2282

2220 TTTTAGTATGAGAATTTGCCCAAGGATACCCCTATGTACTGTAAACAAGAACCCAG 2279
2283 CTGAGGAGCTGTGCTTGGACCAAGCAACTGCCAGTGGAGCCCGGAATCAGGAGTGCAT 2342
2280 CTGAGGAGCTGTGCTTGGACCAAGCAACTGCCAGTGGAGCCCGGAATCAGGAGTGCAT 2339
2343 TGCCCTGCCGAAAAATATCTGTGCAATTTGGCTGGCATTTGGTTGG 2387
2340 TGCCCTGCCGAGTGGCTTGCAGGGTCACTTGGTGTGTGTGGG 2384

RESULT 14
ARI64813
LOCUS ARI64813 2419 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 8 from patent US 6274339.
ACCESSION ARI64813
VERSION ARI64813.1 GI:16238022
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2419)
AUTHORS Moore, K. and Nagle, D. Lynn.
TITILE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 8 14-AUG-2001;
FEATURES
Location/Qualifiers
source 1..2419
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 39.8%; Score 1521.8; DB 6; Length 2419;
Best Local Similarity 80.7%; Pred. No. 18-299;
Matches 1895; Conservative 0; Mismatches 247; Indels 205; Gaps 2;
3 GGTGGCGGACGCGCGCAACTGAGCAAGGCTGAGGAGGAGGAGCGCGCGGACGCGCAGC 62
82 GGTGGCGGCGCGCGCGGCGACTGAGGCGGCTGAGGGGAGGAGCAAGAGACGACGAGC 141
63 GCTGGCGGCGAGAGCGCGCGCGCGCTGAGGAGCTGGGACGTGACCGAGGCTGGAGGCG 122
142 GCCTGGCGGCGAGAGGCGGCGGAGCAGCAGCAGCCTGCAACCGGAGAGGCGCTGGAGG 201
123 GGGGCTGGCGGCGCGGCTGCGCTCCCGGCGCTGCTCTCCACCGCTGGCGGCGCGGCT 182
202 GGGACCGCGCGCGCGCTGCTCTCCCGGCTGCTCTCGGGGCGCTGCGCGGCGCTGCG 261
183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
243 CGAGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
313 CGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
303 TGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
373 TGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
363 CGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
433 CACGGGCTGGTGGGCGAGCAATGCCAGCAGTGGCGGCGCGCTTTCAGGACATCTGCTC 492
411 ----- 410
493 AGCGCTATAATCACAGCTGTTCCGAGGAGTGGAGGAGCAAGTTCGAGGCGAGCTTC 552
411 ----- 410
553 GGCTACAGAAATTAAGTTCAAGAGTAACCTGGGGCAACTTTGGGCTTTGTCTCCAAAACCAAAA 612

411 ----- 410
613 TGACGAAAGGACAGCTAGAGTCTTTTGGGAAATTTTAGCTGACTAATTTTCAAC 672
411 ----- ACTAATCTGATCTTCTGGTTTGTGACAGATGACCTGAAATTTAATAACAAA 466
673 GAGAACTAATCTGCTTCTGATTTGTAAAGATGACCTGGAAATTTAATAATATAAGA 732
467 CGAAGTCACTGCTCTCAATGAAAGGACAGCCAAATAGAAATAATGAGACTTCGTTCAATC 526
733 CGAAGTCACTGCTCTCAATGAAAGGACAGCCAAATAGAAATAATGAGACTTCGTTCAAC 732
527 ATTTTGTACAGATGAGTTGGACCAATTTATATATGTTTATGATGGGACTCAATTTATG 586
793 ATTTTGTACAGATGAGTTGGACCAATTTATATATGTTTATGATGGGACTCAATTTACG 852
587 CACCGTATGCTGCAATTTAGTGGCTCAATGTTCTGAGAGAGATGGCAATGAGACTG 646
853 CACCTCTGATTTGCTGCTTTAGTGGCTCAATGTTCTGAGAGATGGCAATGAGAGCG 912
647 TCCCTGAGGTTGTTGCCACATCAGTTATAGCTTGTGCTGCAATTTTATGATGCTGCTT 706
913 CTCCTGAGGTTGCTGCTTCACTTCACTGCTGCTGCAATTTTATGATGCTGCTT 972
707 ATATTTGACTGGATTTAATATTTACTTACAGTTTGTATGATGTTGCCAAATAGCTCAG 766
973 ATATTTGACTGGATTTAATATTTACTTACAGTTTGTATGATGTTGCCAAATAGCTCAG 1032
767 GCCGAGGAGAGTGAAGATCAGTAATAGACGCAAACTGTTGAAATGTAATGTTCTGAAA 826
1033 GCCGAGGAGAGTGAAGAGCAGTAACAGCAGCAGGCTGTTGAGTGTGAATGTTCTGAAA 1092
827 ACTGAAAGGTGAAGCATGTGACATTTCTTACTGTACAGCACTGTGGTTTCTCTCATC 886
1093 ACTGAAAGGTGAAGCATGTGACATTTCTTACTGTACAGCACTGTGGTTTCTCTCATC 1152
887 GAGCATCTGCAATTTCAAGTGTATGACAGATGCTCTGCTTCTCAGATGCGCAGGCTC 946
1153 GAGCATCTGCAATTTCAAGTGTATGACAGATGCTCTGCTTCTCAGATGCGCAGGCTC 1212
947 CTGATGTTGAGTTCTCTTACAGTAACTGATGATGATGATGATGATGATGATGATGAT 1006
1213 CTGATGTTGAGTTCTCTTACAGTAACTGATGATGATGATGATGATGATGATGATGAT 1272
1007 ACTTAAAGTCTCCAGAGCATCTCAATGAGTGTGTTCAATGAGAAATTTATGAGTTG 1066
1273 ATTTAAAGTCTCCAGAGCATCTCAATGAGTGTGTTCAATGAGAAATTTATGAGTTG 1332
1067 TTGAGGATATATGTTCAACCACTCAGATTTAATGATGATGATGATGATGATGATGATG 1126
1333 TTGAGGATATATGTTCAACCACTCAGATTTAATGATGATGATGATGATGATGATGATG 1392
1127 CTAGGAGTGGCTTCCATTAACCGTCTGTGAACATGTTGTTGATGATGATGATGATGAT 1186
1393 CTAGGAGTGGCTTCCATTAACCGTCTGTGAACATGTTGTTGATGATGATGATGATGAT 1452
1187 CTTTGCATTTACAGGATAAAATTTACATGTTGAGGAAAATTTGATCCTGAGG 1246
1453 CTTTGCATTTACAGGATAAAATTTACATGTTGAGGAAAATTTGATCCTGAGG 1512
1247 ATGTGACCAATGAGTTGAGATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 1306
1513 ACGTGACCAATGAGTTGAGATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 1572
1307 CTAAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
1573 CGAAGCTAAGGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
1367 ATGGCCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426
1633 CTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
1427 ATGTGAGGATATGATTTGGATTAAGAACACATGAGGATATATACACCCAGGTTGCC 1486

1693 TTGTGAGGAATATGACTTTGGAAAGAACACATGAGGATATATATCATCTACTCAGGTGTC 1752
1487 TTGTCAAAGGGGGTTACCGCCATAGCAGTGTTTACGACCATAGGACCGGCGCTTATACG 1546
1753 TTGTCAAAGGGGGTTATGCGCCACAGTGTGTTTATGATGACAGGACCAAGGCTCTGTACG 1812
1547 TTCTATGTTGGCTACAAGGCTTTCACTGCGCAATAGTACCGGCTTCAGATGATCTCTACC 1606
1813 TTCTATGTTGGCTACAAGGCTTTCACTGCGCAATAGTACCGGCTTCAGATGATCTCTACA 1872
1607 GATATGATGATGATACCCAGATGTCGACATTTCTTTAAGGACAGCCGATTTTCCGTTACT 1666
1873 GATACGATGATGATCTCAGATGTCGACATTTCTTTAAGGACAGCCGATTTTCCGTTACT 1932
1667 TCCACACAGCTGTGATGATGATGAGTGAACCATGCTGCTGTTTGGGGGAAACACACAATG 1726
1933 TCCACACAGCTGTGATGATGATGAGTGAACCATGCTGCTGTTTGGGGGAAACACACAATG 1992
1727 ACACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1786
1993 ACATTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2052
1787 CCTGTGACCGCTGCTCAGTGTCTCCAGACCTGATCTCCACATGATGATGATGATGATGAT 1846
2053 CTTGTGACCGATGCTCAGTGTCTCCAGACCTGATCTCCACATGATGATGATGATGATGAT 2112
1847 GCAATTCAGAGCTCTTACACACAGCAGCAGCATGATGATGATGATGATGATGATGATGAT 1906
2113 GCAATTCAGAGCTCTTACACACAGCAGCAGCATGATGATGATGATGATGATGATGATGAT 2172
1907 TCTCTGAGCAGCATCTCTGCTGATTTACCTCGAAGCAGTGTGATGATGATGATGATGATGAT 1966
2173 TCTCTGAGCAGCTCTGCTGATTTACCTCGAAGCAGTGTGATGATGATGATGATGATGAT 2232
1967 CTTGTTTAGCAGAGACCTGCTGATTTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2026
2233 CTTGTTTAGCAGAGACCTGCTGATTTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2292
2027 TCTCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2086
2293 CTTCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2352
2087 AAAGAACTCTTGACCATGACAGATGTGACAGCAGACAGATGTTTACAGTGTGATGATGAT 2146
2353 AAAGAACTCTTGACCATGACAGATGTGACAGCAGACAGATGTTTACAGTGTGATGATGAT 2412
2147 ACACCAA 2153
2413 ATACCAA 2419

RESULT 15
AR164815
LOCUS AR164815
DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815
VERSION AR164815.1 GI:16238024
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6370)
AUTHORS Moore, K. and Nagle, D. Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 12 14-AUG-2001,
FEATURES Location/Qualifiers
 1. 6370
 /organism="unknown"
 /mol_type="unassigned DNA"

Query Match 35.7%; Score 1364.4; DB 6; Length 6370;

Best Local Similarity 95.0%; Pred. No. 1.4e-267; Mismatches 1492; Conservative 7; Indels 23; Gaps 8;

QY	2257	ATGTACTACTGTAAACAAGAACACCACTGCGAGAGCTGTGCGCTCGACAGAACCTGCCAG	2316
DB	1	ATGTACTACTGTAAACAAGAACACCACTGCGAGAGCTGTGCGCTCGACAGAACCTGCCAG	60
QY	2317	TGGAGCCCGGGAATCAGAGTGCATTGCCCTGCCGGAATAATCTGTGGCAATGGCTGG	2376
DB	61	TGGAGCCCGGGAATCAGAGTGCATTGCCCTGCCGGAATAATCTGTGGCAATGGCTGG	120
QY	2377	CAATTTGGTGGAAACTCATGTTTGAATAATTACTGTCCAGGAGAAATTATGACAAATGCT	2436
DB	121	CAATTTGGTGGAAACTCATGTTTGAATAATTACTGTCCAGGAGAAATTATGACAAATGCT	180
QY	2437	AAATTTGCTGTAGGAACCAATGCCCTTTTGGCTTTCTCTTACAAACCAAGAGGTA	2496
DB	181	AAATTTGCTGTAGGAACCAATGCCCTTTTGGCTTTCTCTTACAAACCAAGAGGTA	240
QY	2497	GAATTTGCTTAAAGCAGCTGGAATAATGCAATCATCTCAGAGCATGTCACCTCACC	2556
DB	241	GAATTTGCTTAAAGCAGCTGGAATAATGCAATCATCTCAGAGCATGTCACCTCACC	300
QY	2557	TTAACCCCAATGGCTCGGCTT---CGGAAGATCAATGTGCTCTACTGG---TGCTGG	2608
DB	301	TTAACCCCAATGGCTCGGCTT---CGGAAGATCAATGTGCTCTACTGG---TGCTGG	360
QY	2609	AGATATGTCCCAATTACAAATAG-TTACTACAGTGGATGCCGTCTG---AGCCAG	2663
DB	361	AGATATGTCCCAATTACAAATAG-TTACTACAGTGGATGCCGTCTG---AGCCAG	420
QY	2664	TGATGCT--GGATTTCTGTGGAAATTT---ATCAGAACCCAGTACTCGGGACTGAAGG	2716
DB	421	TGATGCT--GGATTTCTGTGGAAATTT---ATCAGAACCCAGTACTCGGGACTGAAGG	480
QY	2717	CTGCAACTGCAAT--CAACCCACTCAATGTGATGTCTGTGAAAGGCTGCAAAACCAAG	2774
DB	481	CTGCAACTGCAAT--CAACCCACTCAATGTGATGTCTGTGAAAGGCTGCAAAACCAAG	540
QY	2775	TGCTAA--GCAGTCCGGACACCATGCTGCTTTCAGACAGCATGTGGAGATGTCACACCG	2833
DB	541	TGCTAA--GCAGTCCGGACACCATGCTGCTTTCAGACAGCATGTGGAGATGTCACACCG	600
QY	2834	GCAGTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGACTTCCAATGCCATG	2893
DB	601	GCAGTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGACTTCCAATGCCATG	660
QY	2894	TGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTATACATGACGACCTGCCCTCTG	2953
DB	661	TGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTATACATGACGACCTGCCCTCTG	720
QY	2954	AAAATTTGTCAGGCTACTGTACTGTCTGATGCTTGGAGAACCAAGGCTGTGGCTGGT	3013
DB	721	AAAATTTGTCAGGCTACTGTACTGTCTGATGCTTGGAGAACCAAGGCTGTGGCTGGT	780
QY	3014	GTACTGATCCAGCAATFACTGGGAAGGAATGATAGAGGTTCTCTAAGAGGACAG	3073
DB	781	GTACTGATCCAGCAATFACTGGGAAGGAATGATAGAGGTTCTCTAAGAGGACAG	840
QY	3074	TGAAGATGCTTCCGAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTTCCA	3133
DB	841	TGAAGATGCTTCCGAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTTCCA	900
QY	3134	GCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCACTGTCTCAGCTTGCCAAAT	3193
DB	901	GCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCACTGTCTCAGCTTGCCAAAT	960
QY	3194	GCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA	3253
DB	961	GCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA	1020
QY	3254	CAGCAAGCACTGCGAGACCTGCATATCTGGCTTTCTACGGTGTATCCCAATGGAGGA	3313

Search completed: March 5, 2004, 05:11:49
Job time : 9628.23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:59:34 ; Search time 6050.99 Seconds
(without alignments)
18847.119 Million cell updates/sec

Title: US-09-787-097-19
Perfect score: 3819
Sequence: 1 atgggtgcccagcgccggc.....aggtgcaaacatgaacatga 3819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	2940.6	77.0	3976 29 AY418587 - AY418587 Homo sapi
2	2507	65.6	3671 29 AY418588 AY418588 Pan trogl
3	2507	65.6	3976 29 AY418589 AY418589 Mus muscu
4	925.6	24.2	1201 13 BX440935 BX440935

5	775	20.3	964	13	BU506373	AGENCOURT
6	763.2	20.0	933	12	BG678679	BG678679 602624477
7	701.2	18.4	840	14	CB520374	UI-M-GIO-
8	681	17.8	681	12	BM783739	K-EST0061
9	680.4	17.8	683	12	BM783788	K-EST0061
10	680	17.8	680	12	BM783258	K-EST0061
11	680	17.8	680	12	BM783773	K-EST0061
12	677	17.7	677	12	BM783718	K-EST0061
13	674.4	17.7	677	12	BM783821	K-EST0061
14	673	17.6	673	12	BM783685	K-EST0061
15	669.8	17.5	962	13	BX431593	BX431593
16	663.8	17.4	792	14	CD644584	AGENCOURT
17	636.8	16.7	674	12	BG563729	BG563729 602584512
18	636.4	16.7	770	13	BY761276	BY761276 602584512
19	636	16.7	777	14	CF745736	CF745736 602584512
20	632.4	16.6	706	12	BI089332	602584512
21	632	16.5	632	12	BM773003	K-EST0057
22	632	16.5	632	12	BM783264	K-EST0061
23	626.6	16.4	953	10	BF313142	601897661
24	623	16.3	917	10	BF672370	602150605
25	618	16.2	633	10	BE894518	601433126
26	610.6	16.0	722	13	BX669806	BX669806
27	608	15.9	735	13	BY761390	BY761390
28	607.8	15.9	991	10	BE880905	601490448
29	597.6	15.6	940	13	BQ956223	AGENCOURT
30	585.2	15.3	716	14	CF745901	UI-M-BH2-
31	580.2	15.2	714	10	B5650220	UI-M-BH2-
32	580	15.2	591	12	BI850289	imageqc_1
33	568.2	14.9	653	14	CF446934	CF446934 700912 MA
34	554.2	14.5	663	14	CF745835	CF745835 UI-M-GVO-
35	548	14.3	902	13	BU151781	AGENCOURT
36	537.4	14.1	644	14	CB425680	600762 MA
37	535.8	14.0	942	10	BF316716	601903293
38	535.2	14.0	595	10	BF853681	MR2-EN009
39	534.2	14.0	598	10	BF853006	MR2-EN009
40	530.6	13.9	641	10	BB664616	BB664616
41	527.2	13.8	544	10	BF853962	MR2-EN009
42	524.2	13.7	626	10	BE980188	UI-M-BG2-
43	521	13.6	600	14	CB432517	608730 MA
44	518	13.6	746	14	CB242644	UI-CF-FNC
45	508.6	13.3	560	10	BF853965	MR2-EN009

ALIGNMENTS

3976 bp DNA linear GSS 17-DEC-2003

AY418587 Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION AY418587
VERSION AY418587.1 GI:39774547
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3976)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3976)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission

TITLE

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source 1..3976
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3976
/gene="ATRN"
/locus_tag="HCM6607"

gene

ORIGIN

Query Match 77.0%; Score 2940.6; DB 29; Length 3976;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2943; Conservative 0; Mismatches 546; Indels 0; Gaps 0;

QY 315 TGTCAACGGCGTGTGTCACACCTGGCACCGGCGAGTGCCTGCCCCGGCGGCTGGGT 374
Db 1 TGTCAACGGCGTGTGTCACACCTGGCACCGGCGAGTGCCTGCCCCGGCGGCTGGGT 60

QY 375 GGGCGAGCAATGCCAGCACTGGGCGGCGCTTCAGACTAAGTGTCTTCTGGGTTGT 434
Db 61 GGGCGAGCAATGCCAGCACTGGGCGGCGCTTCAGAGNNNNNNNNNNNNNNNNNN 120

QY 435 GACAGATGACCTGGAAATATATAAATACAAACGAAGTGCAGTGGCTCAITGAAGGACA 494
Db 121 NNN 180

QY 495 GCCAATAGAAATAGAGACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTTGGACCA 554
Db 181 GCCAATAGAAATAGAGACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTTGGACCA 240

QY 555 TTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGCT 614
Db 241 TTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGNNNN 300

QY 615 CATTGCTCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTGCCACATCAGGTTA 674
Db 301 NNN 360

QY 675 TGCCCTTGCTGCATTTTITAGTGCTGCTTAAATTTTGACTGGAATTAATATCTTA 734
Db 361 NNN 420

QY 735 CAGTTTTCATATGTTCCAAATAACTGCTCAGGCGGAGAGTGTAAAGATCAGTAATAG 794
Db 421 NNNTTTTCATATGTTCCAAATAACTGCTCAGGCGGAGAGTGTAAAGATCAGTAATAG 480

QY 795 CAGCGAACTGTTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGATGTGACATCC 854
Db 481 CAGCGAATCTGTTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGATGTGACATCC 540

QY 855 TCACGTACAGACAACTGTGGTTTTCCTCATCGAGGATCTGCAATTCAGTGTATGTCAG 914
Db 541 TCACGTACAGACAACTGTGGTTTTCCTCATCGAGGATCTGCAATTCAGTGTATGTCAG 600

QY 915 AGGATGCTCTGCTTCTCAGACTGGCGGGTCTGGATGTTTCAGTCTGTACAGCTAA 974
Db 601 AGGATGCTCTGCTTCTCAGACTGGCGGGTCTGGATGTTTCAGTCTGTACAGCTAA 660

QY 975 CCAGTCATTTTGGACTCAGAGGAATATTTCTAATTAAGCTCCCGAGAGCATCTCATAA 1034
Db 661 CCAGTCATTTTGGACTCAGAGGAATATTTCTAATTAAGCTCCCGAGAGCATCTCATAA 720

QY 1035 AGCTGTGCTCAATGAAACATTATGTGGTTGTTGGAGATATATGTTCAACCACTCAGA 1094
Db 721 AGCTGTGCTCAATGAAACATTATGTGGTTGTTGGAGATATATGTTCAACCACTCAGA 780

QY 1095 TTATAACATGTTCTAGGATGACCTTGTCTTCTAGGAGTGGCTTCCACTAACCGTTC 1154
Db 781 TTATAACATGTTCTAGGATGACCTTGTCTTCTAGGAGTGGCTTCCACTAACCGTTC 840

QY 1155 TGTCAACAAATGTGGTTGTTAGATATGTCATCTCTTTGGCAATTAACAAGGATAAAATTTA 1214
Db 841 TGTCAACAAATGTGGTTGTTAGATATGTCATCTCTTTGGCAATTAACAAGGATAAAATTTA 900

QY 1215 CATGTATGGAGGAAATTCATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTC 1274
Db 901 NNN 960

QY 1275 CATTCAATAATGAGTCAATGGGTGTGTGACCCCTAAGCAAGGAGAGCATGTCAGTGT 1334
Db 961 NNN 1020

QY 1335 TGGGCACTCTGCACACATCTTACATCTGAAGATGCCGAGTGGTCACTGCTGTCATCTT 1394
Db 1021 NNN 1080

QY 1395 TGTCTCATCTCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAATGTTTGGATAAGAA 1454
Db 1081 NNN 1140

QY 1455 CACATGGAGTATATTACACACCCAGGTCCTTGTGCAAGGGGTTACGGCCATAGCAG 1514
Db 1141 CACATGGAGTATATTACACACCCAGGTCCTTGTGCAAGGGGTTACGGCCATAGCAG 1200

QY 1515 TGTTTACGACCATAGACACGAGGCGCTTATACGTTTCACTGTGGCTACAAGGCTTTCAGTGC 1574
Db 1201 TGTTTACGACCATAGACACGAGGCGCTTATACGTTTCACTGTGGCTACAAGGCTTTCAGTGC 1260

QY 1575 CAATAAGTACCGGCTTCAGATGATCTCTACCGATATGATGGAATACCCAGATGAGAC 1634
Db 1261 CAATAAGTACCGGCTTCAGATGATCTCTACCGATATGATGGAATACCCAGATGAGAC 1320

QY 1635 CATTCTTAAGGACAGCGGATTTTCCGTTACTTGGCACAGCTGTGATAGTACAGTGGAGAC 1694
Db 1321 CATTCTTAAGGACAGCGGATTTTCCGTTACTTGGCACAGCTGTGATAGTACAGTGGAGAC 1380

QY 1695 CATGCTGTGTTTGGGGGAAACACACAAATGACATCTATGAGCCATGGGCCCAATG 1754
Db 1381 CATGCTGTGTTTGGGGGAAACACACAAATGACATCTATGAGCCATGGGCCCAATG 1440

QY 1755 CTTCTCTCAGATTTTATGCGCTATGACATTCCTGTGACCGCTGGTCACTGCTCCAG 1814
Db 1441 CTTCTCTCAGATTTTATGCGCTATGACATTCCTGTGACCGCTGGTCACTGCTCCAG 1500

QY 1815 ACCTGATCTCCACCATGATGTCAACAGATTTTGGCCATTCAGCAGTCTTACACACAGCAC 1874
Db 1501 NNN 1560

QY 1875 CAGTATGTTTGGTGGTTTCAATAGTCTCTCTCAGCGGATCTCTGGGTATTCACCTC 1934
Db 1561 CATGATGTTTGGTGGTTTCAATAGTCTCTCTCAGCGGATCTCTGGGTATTCACCTC 1620

QY 1935 GGAACAGTGTGATGCGCATCGGAGTGAAGCCCTTGTGTAGCAGCAGACCTGGTATTCG 1994
Db 1621 GGAACAGTGTGATGCGCATCGGAGTGAAGCCCTTGTGTAGCAGCAGACCTGGTATTCG 1680

QY 1995 GTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTGTGGGCGCTGGCAACTGTATGAACA 2054
Db 1681 GTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTGTGGGCGCTGGCAACTGTATGAACA 1740

QY 2055 AGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTA 2114
Db 1741 AGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTA 1800

QY 2115 CCAGCACACAGATTTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGAATGA 2174
Db 1801 CCAGCACACAGATTTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGAATGA 1860

QY 2175 CCATTGTGTCACAGGACCAACAGCTGTGACAGGCGGATCTCCATTTTATAGTATGA 2234
Db 1861 CCATTGTGTCACAGGACCAACAGCTGTGACAGGCGGATCTCCATTTTATAGTATGA 1920

QY 2235 GAATGCCCCAAGGATTAACCCCATGTACTACTGTAAACAAGACACAGCTGCGAGGCTG 2294
Db 1921 GAATGCCCCAAGGATTAACCCCATGTACTACTGTAAACAAGACACAGCTGCGAGGCTG 1980
QY 2295 TGCCCTGAGACAGAACTGCGAGTGGGAGCCCGGAATCAGAGTGCATTCGCCCTGCCCGA 2354
Db 1981 TGCCCTGAGACAGAACTGCGAGTGGGAGCCCGGAATCAGAGTGCATTCGCCCTGCCCGA 2040
QY 2355 AAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACTCATGTTGAAAATTAATCTGTC 2414
Db 2041 AAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACTCATGTTGAAAATTAATCTGTC 2100
QY 2415 CAAGAGAAATTAACAATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTC 2474
Db 2101 CAAGAGAAATTAACAATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTC 2160
QY 2475 TCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAATAATGCAATC 2534
Db 2161 TCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAATAATGCAATC 2220
QY 2535 TCAGAGCATGTCAGCTCAACCTTAACCCCATGGCTCGGCTTCGGAAGATCAATGTGTC 2594
Db 2221 TCAGAGCATGTCAGCTCAACCTTAACCCCATGGCTCGGCTTCGGAAGATCAATGTGTC 2280
QY 2595 CTACTGTTGCTGGGAAGATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2654
Db 2281 CTACTGTTGCTGGGAAGATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2340
QY 2655 TGAGCCCACTGATCTGGATTTCTGGAAATTTATCAGAACCCAGTACTCGGGAGCTGAA 2714
Db 2341 TGAGCCCACTGATCTGGATTTCTGGAAATTTATCAGAACCCAGTACTCGGGAGCTGAA 2400
QY 2715 GGCTGCAACCTGATCAACCCATCAATGTTAGTGTCTGTGAAAGGCTGCAACACACAG 2774
Db 2401 GGCTGCAACCTGATCAACCCATCAATGTTAGTGTCTGTGAAAGGCTGCAACACACAG 2460
QY 2775 TGCTAAGCAGTGCAGGACACCATGTCCTTGAGACAGCATGTGGAGATTGCACACGCG 2834
Db 2461 TGCTAAGCAGTGCAGGACACCATGTCCTTGAGACAGCATGTGGAGATTGCACACGCG 2520
QY 2835 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAAGCAGTGTGGACCTCCATGCTATGT 2894
Db 2521 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAAGCAGTGTGGACCTCCATGCTATGT 2580
QY 2895 GGCTCTCTCTCTTTGGCCAGTGTATGGAATGTTATACGATGAGCAGCTGCCCTGTA 2954
Db 2581 GGCTCTCTCTCTTTGGCCAGTGTATGGAATGTTATACGATGAGCAGCTGCCCTGTA 2640
QY 2955 AAATGTTTCAGGCTACTGTACTGTAGTCACTGTGAGCAACACAGGCTGTGGCTGGTG 3014
Db 2641 AAATGTTTCAGGCTACTGTACTGTAGTCACTGTGAGCAACACAGGCTGTGGCTGGTG 2700
QY 3015 TACTGATCCAGCAATGCTGCAAGGGAATGATGATGAGGGTTCTTAAAGACACAGT 3074
Db 2701 TACTGATCCAGCAATGCTGCAAGGGAATGATGATGAGGGTTCTTAAAGACACAGT 2760
QY 3075 GAAGATGCTTTGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCAG 3134
Db 2761 GAAGATGCTTTGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCAG 2820
QY 3135 CATGTGCTAGAGGACAGATCAACTGCTTTTCAATTCATGTCAGCTTCCCAATG 3194
Db 2821 CATGTGCTAGAGGACAGATCAACTGCTTTTCAATTCATGTCAGCTTCCCAATG 2880
QY 3195 CAAGGCGCACAGTAATCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACAC 3254
Db 2881 CAAGGCGCACAGTAATCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACAC 2940
QY 3255 AGGCAAGCACTGCGAGACCTGATATCTGGCTTTTACGGTATCCCAATGAGGGAA 3314
Db 2941 AGGCAAGCACTGCGAGACCTGATATCTGGCTTTTACGGTATCCCAATGAGGGAA 3000
QY 3315 ATGTGAGCCATGCAAGTGCATGGGACCGCTCTCTGTGCAACCAACACGCGCAAGTG 3374

RESULT 2

AY418588

LOCUS

DEFINITION

Pan troglodytes ATRN gene, 3671 bp DNA linear GSS 17-DEC-2003

genomic survey sequence.

ACCESSION

AY418588

VERSION

AY418588.1

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

REFERENCE

1 (bases 1 to 3671)

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.

TITLE

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

JOURNAL

Science 302 (5652), 1960-1963 (2003)

PUBMED

14671302

REFERENCE

2 (bases 1 to 3671)

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.

TITLE

Direct Submission

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT

These sequences were made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Location/Qualifiers

1..3671

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

<1..>3671

gene


```

/ gene = "ATRN"
/ locus_tag = "HCM6607"

ORIGIN
Query Match          65.6%; Score 2507; DB 29; Length 3671;
Best Local Similarity 76.0%; Pred. No. 0;
Matches 2516; Conservative 0; Mismatches 793; Indels 0; Gaps 0;

QY 495 GCCAAATAGAATAATGAGACTTCGTTTCAATCAITTTTGCTACAGAGTGAGTTGGGACCA 554
    |||||
Db 1 GCCAAATAGAATAATGAGACTTCGTTTCAATCAITTTTGCTACAGAGTGAGTTGGGACCA 60
    |||||
QY 555 TTTATATGTTTTATGATGGGACTCAATTTATGACCCGCTAGTGTGCTGCATTTAGTGGCCT 614
    |||||
Db 61 TTTATATGTTTTATGATGGGACTCAATTTATGACCCGCTAGTGTGCTGCATTTAGNNNNN 120
    |||||
QY 615 CATTGTTCCCTGAGAGAGATGCCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGGTTA 674
    |||||
Db 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
    |||||
QY 675 TGCCTTGCTGCATTTTTTAGTGATGCTGCTTATAA TTTGACTGGA TTTAA TATACTTTA 734
    |||||
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
    |||||
QY 735 CAGTTTTGATATGTTGCCAATAACTGCTCAGGCCGAGGAGTGTAAGATCAGTAATAG 794
    |||||
Db 241 NNNTTTGGTATGTTGCCAATAACTGCTCAGGCCGAGGAGTGTAAGATCAGTAATAG 300
    |||||
QY 795 CAGCGAAACTGTTGAATGTGAATGTTCTGAAAAC TGGAAAGGTGAAGCATGTGCATTC 854
    |||||
Db 301 CAGCGAATACTGTTGAATGTGAATGTTCTGAAAAC TGGAAAGGTGAAGCATGTGCATTC 360
    |||||
QY 855 TCACGTGTACAGACAAC TCGTGGTTTTTCTCATCGAGGCATCTGCAATTTCAAGTGATGTCAG 914
    |||||
Db 361 TCACGTGTACAGACAAC TCGTGGTTTTTCTCATCGAGGCATCTGCAATTTCAAGTGATGTCAG 420
    |||||
QY 915 AGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTTGGATGTTTCAGTTCCTGTACACAGCTAA 974
    |||||
Db 421 AGGATGCTCCTGCTTCTCAGACTGGCAGGNNCTGGATGTTTCAGTTCCTGTACACAGCTAA 480
    |||||
QY 975 CCAGTCATTTTGGACTCGAGAGGAATA TTTCTAACTTAAAGTCCCGAGAGCATCTATAA 1034
    |||||
Db 481 CCAGTCATTTTGGACTCGAGAGGAATA TTTCTAACTTAAAGTCCCGAGAGCATCTATAA 540
    |||||
QY 1035 AGCTGTGGTCAA TGGAAACA TTAGTGGGTGTTGGAGGATATATGTTCAACCACTCAGA 1094
    |||||
Db 541 AGCTGTGGTCAA TGGAAACA TTAGTGGGTGTTGGAGGATATATGTTCAACCACTCAGA 600
    |||||
QY 1095 TTATAACATGGTCTACGGTATGACCTTGCTTTCTAGGGAGTGGCTTCCACTAAACCGTTC 1154
    |||||
Db 601 TTATAACATGGTCTACGGTATGACCTTGCTTTCTAGGGAGTGGCTTCCACTAAACCGTTC 660
    |||||
QY 1155 TGTGCAACAATGTTGTTGTAGATATGTCATCTTTTGCACTATATACAAGGATATAAATTTA 1214
    |||||
Db 661 TGTGCAACAATGTTGTTGTAGATATGTCATCTTTTGCACTATATACAAGNNNNNNNNNN 720
    |||||
QY 1215 CATGTATGGAGGAAAA TTAGTCAACTGGGAATGTGCCAATGAGTTGAGAGTTTTCAT 1274
    |||||
Db 721 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
    |||||
QY 1275 CATTCATAATGAGTCATGGGTGTTGTGACCCCTAAGGCAAGGAGCAGTATGCAGTGGT 1334
    |||||
Db 781 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 840
    |||||
QY 1335 TGGGCAC TCTGCACACATTTGTCAC TGAAGAA TGGCCGAGTGGTCATGCTGTCATCTT 1394
    |||||
Db 841 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 900
    |||||
QY 1395 TGGTCAC TCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAGAA 1454
    |||||
Db 901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960
    |||||
QY 1455 CACATGGAGTATATTA CACACCCAGGGTGCCTTGTGCAAGGGGGTTACCGGCATAGCAG 1514
    |||||

```

[illegible]

2041 TCNNNNNNNTCCAAAGCTCACCTTAACCCCATGGGTGGGCTCGGAGATCAATGTGTC 2100
2595 CTACTGCTGCTGGGAAGATATGTCCTCCCATTTTACAAATAGTTTACTACAGTGGATGCGGTC 2654
2101 CTACTGCTGCTGGGAAGATATGTCCTCCCATTTTACAAATAGTTTACTACAGTGGATGCGGTC 2160
2655 TGAGCCCAAGTATGCTGGATCTGTGGAAATTTTATCAGAACCCAGTATCTCGGGACTGAA 2714
2161 TGAGCCCAAGTATGCTGGATCTGTGGAAATTTTATCAGAACCCAGTATCTCGGGACTGAA 2220
2715 GGCCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTGCAAAACCCAG 2774
2221 GGCCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTGCAAAACCCAG 2280
2775 TGCTAAGCAGTCCCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTCACCCAGCGG 2834
2281 TGCTAAGCAGTCCCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTCACCCAGCGG 2340
2835 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGGACTCCAATGCTATGT 2894
2341 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGGACTCCAATGCTATGT 2400
2895 GGCCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGTGTGCCCCCTGA 2954
2401 GGCCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGTGTGCCCCCN 2460
2955 AAATGTTCAAGCTACTGCTACTGCTAGTCTGTTGGAGCAACAGGCTGTGGCTGTG 3014
2461 NNN 2520
3015 TACTGATCCAGCAATCTGCGCAAGGAAATGCATAGAGGTTCTCTATAAGGACCAGT 3074
2521 NNN 2580
3075 GAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATCCAG 3134
2581 NNN 2640
3135 CATGCTGTAGAGGACAGAGATACAACTGGTCTTCTTCTACTCTCAGCTGCGCAATG 3194
2641 NNN 2700
3195 CAAGGCGCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTCAGCAC 3254
2701 CAAGGCGCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTCAGCAC 2760
3255 AGGCAAGCATCTCGAGACCTGATCTGGCTTCTACGGTGTATCCCAATGGAGGAA 3314
2761 AGGCAAGCATCTCGAGACCTGATCTGGCTTCTACGGTGTATCCCAATGGAGGAA 2820
3315 ATGTCAGCCATCAAGTGAATGGGACGCGTCTGTGCAACACCAACAGCGGCAAGTG 3374
2821 NNN 2880
3375 TTCTGCAACCAAGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAAATCG 3434
2881 TTCTGCAACCAAGGGCGTCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2940
3435 ATACCAAGAAACCTCTCAGAGGAACATGTTATATCTCTCTTATGATGATCAGTT 3494
2941 ATACNN 3000
3495 CACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTTGTGGCTAC 3554
3001 CACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTTGTGGCTAC 3060
3555 TCCTGACGAACAAAACAGGATTTGGACATGTTTCAATCAATGCTTCAGAAATTTCAACT 3614
3061 TCCTGACGAACAAAACAGGATTTGGACATGTTTCAATCAATGCTTCAGAAATTTCAACT 3120
3615 CACATCACTGGCTGCGAGTTTCTCAGCTGCAACCCAGGCTGGAGAGAGATGCTGT 3674
3121 CACATCACTGGCTGCGAGTTTCTCAGCTGCAACCCAGGCTGGAGAGAGATGCTGT 3180

3675 TGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTTCTCTAAATGAGAAGTTGATT 3734
3181 TGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTTCTCTAAATGAGAAGTTGATT 3240
3735 TCGCAACCCACCAAAATATCACTTCTTGTATGTAGTATGTCAGTAATTCACCTGGCCCAATCAA 3794
3241 TCGCAACCCACCAAAATATCACTTCTTGTATGTAGTATGTCAGTAATTCACCTGGCCCAATCAA 3300
3795 AATTCAGT 3803
3301 AATTCAGT 3309

RESULT 3
AY418589
LOCUS
DEFINITION
Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY418589
VERSION
AY418589.1 GI:39774549
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 3976)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED
14671302
REFERENCE
2 (bases 1 to 3976)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
COMMENT
Location/Qualifiers
FEATURES
source
1..3976
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..3976
/gene="ATRN"
/locus_tag="HCM6607"

ORIGIN
Query Match 65.6%; Score 2507; DB 29; Length 3976;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 817; Indels 0; Gaps 0;
Qy 315 TGTCAAGCGGTCGCTGCAACCTGCGCAGCGGCGCTTCAGACTAATCGATCTTCTGGGTTGT 434
Db 1 TGTCAAGCGGCGCGCTGCAACCTGCGCAGCGGCGCTTCGCCCCACGGCTGGT 60
Qy 375 GGGCGAGCAATGCGCAGCACTGCGGGGCGCTTCAGACTAATCGATCTTCTGGGTTGT 434
Db 61 GGGCGAGCAATGCGCAGCACTGCGGGGCGCTTCAGNNNNNNNNNNNNNNNNNNNN 120
Qy 435 GACAGATGGAGCTGGAATTAATAACAAACGAGTGCAGCTGGCTCATTTGAGACGA 494
Db 121 NNN 180
Qy 495 GCCAAATAGATAATAGAGCTTCGTTTCAATCATTTGCTACAGAGTGTAGTGGACCA 554

Db 181 GCCAAATAGAAATATGAGACTTCGGCTTCAACCATTTTGTCTACAGAAATGTAGCTGGGACCA 240
Qy 555 TTATATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGTCTGCTATTTAGTGGCT 614
Db 241 TTATATATGTTTATGATGGGACTCAATTTACGACCTCTGATTTGCTGCTTTAGNNNNN 300
Qy 615 CATTTGCTCTGAGAGAGATGGCAATCAGACTGTCCTGAGTTGTGCCACATCAGTTA 674
Db 301 NNN 360
Qy 675 TGCCTTGTCTGCAATTTTATTTAGTGATGCTCTTATAATTTGACTGCAATTTAATTTACTTA 734
Db 361 NNN 420
Qy 735 CAGTTTGTATGTGTCCAAATACTGCTCAGCGCGAGAGAGTGTAAAGTACAGTAATAG 794
Db 421 NNNTTTTCATGTGTCCGAATAATTTGCTCAGCGCGAGAGAGTGTAAAGCAGTAAACAG 480
Qy 795 CAGCGAAACTGTTGAATGTGAATGTTCTGAAACTTGGAAAGTGAAGCATGTGCAATTC 854
Db 481 CAGCAGCGCTGTTGAGTGTGAATGTTCTGAAACTTGGAAAGTGAAGCATGTGCAATTC 540
Qy 855 TCACTGTACAGAACTGTTGTTTCTCATCAGAGCATCTGCAATTCAGTGTAGTGTGAG 914
Db 541 TCACTGTACAGAACTGTTGTTTCTTCACTCAGCGAGCATCTGTAATGCAAGCGATACCAG 600
Qy 915 AGGATGCTCTGCTTCTCAGACTGCGAGGCTCTGAGTGTTCAGTTCTGTTACAGCTAA 974
Db 601 AGGATGCTCTGCTTCTTCTCACTGCGAGGCTCTGAGTGTTCAGTTCTGTTACAGCTAA 660
Qy 975 CCAGTCAATTTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAGAGCATCTCATPA 1034
Db 661 CCAGTCTTTTGGACTCGAGAGCAATTTCTGATTTAAAGCTCCCGAGAGCATCTCATPA 720
Qy 1035 AGCTGTGTCATAGAAACATTTATGTGGTGTGTGGAGATATATGTTCAACCACTCAGA 1094
Db 721 AGCTGTGTCATAGAAACATTTATGTGGTGTGTGGAGATATATGTTCAACCACTCAGA 780
Qy 1095 TTATAACATGTTCTAGCGTATGACTTCTTCTAGGAGTGGCTTCCACTTAAACCGTTC 1154
Db 781 TTACAGCATGTTTGTAGCGTATGACTGCTTCTAGGAGTGGCTTCCACTTAAACCGTTC 840
Qy 1155 TGTGAACATGTTGTTGTATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1214
Db 841 TGTGAACATGTTGTTGTGAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 900
Qy 1215 CATGTATGAGGAGAAATTTGATCCAACTGGGAAATGTGACCAATGATGTTGAGATTTTCA 1274
Db 901 NNN 960
Qy 1275 CATTCATAATGAGTATGAGTGTGTTGACCCCTTAAGCAAGAGCAGTATGCAAGTGT 1334
Db 961 NNN 1020
Qy 1335 TGGGCACTCTGACACATGTTTACACTGAAGATGGCCGAGTGTGCTGCTGCTCATCTT 1394
Db 1021 NNN 1080
Qy 1395 TGGTCACTGCCCTCTATGATATATAAGCAATGTGAGCAATATGATTTGGATAAGAA 1454
Db 1081 NNN 1140
Qy 1455 CACATGGAGTATTTATACACCCAGGGTGCCTTGTGCAAGGGGTGTACGGCCATAGCAG 1514
Db 1141 CACATGGAGTATTTATACATCTCAGGGTGTCTTGTGCAAGGGGTGTATGGCCACAGTAG 1200
Qy 1515 TGTTTACGACCAATAGGACACAGGGCTTATAGCTTCTGTTGCTTACAGGGTCTTCAAGTGC 1574
Db 1201 TGTTTATGATGACAGCAAGGCTCTGTGATGTTCAITGGTGTGCTTACAGGGCTTTCAGCG 1260
Qy 1575 CAATAAGTACCGGCTTGCAGATGATCTCTCAACGATGATGTGATACCCAGATGTGAC 1634
Db 1261 CAACAAATACCGGCTTGCAGATGATCTCTCAACGATGATGTGATGATGTGATGTGATGTGAC 1320

Qy 1635 CATTTTAAAGACAGCGGATTTTTCGTTACTTTCACACAGCTGTGATAGTGTGGAAC 1694
Db 1321 CATTTTAAAGACAGCGGATTTTTCGTTACTTTCACACAGCTGTGATAGTGTGGAAC 1380
Qy 1695 CATGCTGTTTGGGGGAAACACACATGATGACATCTATGAGCATGGGCCAAATG 1754
Db 1381 CATGCTGTTTGGAGGAAACACACATGATGACATCTTCCATGAGCCAGGTCGCAATG 1440
Qy 1755 CTTCTCTTCAGATTTTATGCGCTTATGACATGCTGTGACCGCTGTGCTGCTCCAC 1814
Db 1441 CTTCTCTTCAGATTTTATGCGCTTATGACATGCTGNNNNNNNNNNNNNNNNNNNNNNNNNN 1500
Qy 1815 ACCTGATCTCCACATGATGTCAACAGATTTGCGCATTCAGCAGTCTTACACACAGCAC 1874
Db 1501 NNN 1560
Qy 1875 CATGTATGTTTCGCTGGTTCATATAGTCTCTCTCAGCAGCATCTCTGTATTCACCTC 1934
Db 1561 CATGTATGTTTCGCTGGTTCATCAGCTCTCTCTCAGTGTGCTCTGCTTTTACCTC 1620
Qy 1935 GGAACAGTGTGATGCGCATGCGAGTGAAGCGCTTGTAGCAGCAGGACCTGTGTTTCG 1994
Db 1621 GGAACAGTGTGATGCGCATGCGAGTGAAGCGCTTGTGCGCAGCAGGACCTGTGTTTCG 1680
Qy 1995 GTGTGTGTGAAACACAGGCTGCTCAGTGTATCTCTGCGCGCTGGCAACTGATGAACA 2054
Db 1681 GTGTGTGTGAAACACAGTGTCTGATGATCTCTGCGAGTGGCAACTGATGAACA 1740
Qy 2055 AGAAGAAAGTTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTA 2114
Db 1741 AGCAGAAAGTTTAAATCAGAGTGTGTTTTTAAAGAACTTGAACCATGACAGATGTA 1800
Qy 2115 CCAGCACAGATGTTGATGCTGTACAGCCACCAATGACTGCTGCTGTGCAATGA 2174
Db 1801 CCAGCACAGATGTTGATGCTGTACAGCCACCAATGACTGCTGCTGTGCAATGA 1860
Qy 2175 CCATGTTGCTCCCGAGAAACACAGCTGCTCAGAAAGCCAGATCTCCATTTTATGATGA 2234
Db 1861 TCATGTTGCTCCCGAGTGTGAACACAGCTGTACAGAGCCAGATCTCCATTTTATGATGA 1920
Qy 2235 GAATGCTCCCGAGGATTAACCCATGATCTACTGTATACAGAGACAGCTGTGAGAGCTG 2294
Db 1921 GAGTTGCTCCCGAGGATTAACCCATGATCTACTGTATACAGAGAAACAGCTGTGAGAGCTG 1980
Qy 2295 TGCCCTGGAACCACTGCGAGTGGGAGCCCGGAATCAGAGTGTGATGCTGCTGCCGA 2354
Db 1981 TGCCCTGAGACCACTGCGAGTGGGAGCCCGGAATCAGAGTGTGATGCTGCTGCCGA 2040
Qy 2355 AAATATCTGTGGCATTTGGTGTGATTTGTTGAAACTCATGTTTGAATAATTTACTATGC 2414
Db 2041 AAATATCTGTGGCAATTTGGTGTGATTTGTTGAAACTCATGTTTGAATAATTTACTATGC 2100
Qy 2415 CAAGGAGATTTATGACATGCTTAAATGTTCTGTAGAACCAATGCTGCTTTTGGCTTC 2474
Db 2101 TAAGGAGATTTATGACATGCTTAAATGTTCTGTAGAACCAATGCTGCTTTTGGCTTC 2160
Qy 2475 TCTTACAAACCCAGAGAGTGAATTTTCTTAAAGAGCTGCGAATAATCAGTGTATC 2534
Db 2161 CTTCAATCCAGAGAGTGAATTTTCTTAAAGAGCTGCGAATAATCAGTGTATC 2220
Qy 2535 TCAGACATGTCAGAGCTCACCTTAACCCCATGGGTTCGGAGTCAATGTGTC 2594
Db 2221 TCAAGATGTCAGAGCTCACCTTAACCCCATGGGTTCGGAGTCAATGTGTC 2280
Qy 2595 CTACTGCTGCGAGAGATATGTCCTCCCAATTTTAACTTACTACAGTGTGATGCCCTC 2654
Db 2281 TTACTGCTGCGAGAGATATGTCCTCATTTCAATAATGTTGCTGAGTGTGATGCCATC 2340
Qy 2655 TGAGCCAGTGTGATGCTGAGTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGAGTGA 2714
Db 2341 TGAGCCAGTGTGATGCTGAGTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGAGTGA 2400

QY	2715	GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTGCAAAACACAG	2774
Db	2401	GGCTGCAACCTGCATCAACCCCTCTCAATGGCAGCGTCTGTGAAGGCGCTGCAAAACACAG	2460
QY	2775	TGCTAAGCAGTGCAGGACACCATGTCCTTGACAGCAGCATGTGGAGATTGCACACGCG	2834
Db	2461	TGCCAAGCAGTGCAGGACACCATGTCCTTGACAGCAGCATGTGGAGATTGCACACGCG	2520
QY	2835	CAGCTCTGAGTGCATGTGGTGCAGCAACATGAACAGTGTGTGACATCCAAATGCTTATGT	2894
Db	2521	CAGCTCGAGTGCATGTGGTGCAGTACATGAACAGTGTGTGACATCCAAATGCTTACGT	2580
QY	2895	GGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACGATGAGCAGCTGCCCCCTGTA	2954
Db	2581	GGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACGATGAGCAGCTGCCCCCTGTA	2640
QY	2955	AAATGTTTCAAGTCTACTCTGCTAGTCTGTTGAGCAACAGCGTGTGCTGGTG	3014
Db	2641	AAATGTTTCAAGTCTACTCTGCTAGTCTGTTGAGCAACAGCGTGTGCTGGTG	2700
QY	3015	TACTGATCCCAAGCAATCTGCAAGGGAATGATAGAGGGTTCCTATAAGACCACT	3074
Db	2701	TACTGATCCCAAGCAATCTGCAAGGGAATGATAGAGGGTTCCTATAAGACCACT	2760
QY	3075	GAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATTCAG	3134
Db	2761	GAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTTCTGAATCCAG	2820
QY	3135	CATGTGCTTAGAGGACAGACATCAACTGGTCTTTTCAATTCAGTCTGCTGCTTCCCAATG	3194
Db	2821	CATGTGCTTAGAGGACAGACATCAACTGGTCTTTTCAATTCAGTCTGCTGCTTCCCAATG	2880
QY	3195	CAACGGCCACAGTAAATGCATCAATCAGACATCTGTGAGAAAGTGTGAGAACCTGACCC	3254
Db	2881	CAACGGCCACAGTAAATGCATCAATCAGACATCTGTGAGAAAGTGTGAGAACCTGACCC	2940
QY	3255	AGGCAAGCACTCGAGACCTGATATCTGCTTCTACCGTATCCCAATGAGGAGAA	3314
Db	2941	AGGCAAGCACTCGAGACCTGATATCTGCTTCTACCGTATCCCAATGAGGAGAA	3000
QY	3315	ATGTCAGCCATGCAAGTGCATGGCCAGCGCTCTCTGTGCAACCAACACAGGCAAGTG	3374
Db	3001	ATGTCAGCCATGCAAGTGCATGGCCAGCGCTCTCTGTGCAACCAACACAGGCAAGTG	3060
QY	3375	CTTCTGCAACCAAGGCGGTCAAGGGGAGAGTGCAGTATGTGAGGTAGAAATCG	3434
Db	3061	CTTCTGCAACCAAGGCGGTCAAGGGGAGAGTGCAGTATGTGAGGTAGAAATCG	3120
QY	3435	ATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTTCTTATTGACTATCAGTT	3494
Db	3121	ATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTTCTTATTGACTATCAGTT	3180
QY	3495	CACCTTTAGTCTATCCAGGAAGATGCTGCTATTATACAGCTATCAATTTTGGCTAC	3554
Db	3181	CACCTTTAGTCTATCCAGGAAGATGCTGCTATTATACAGCTATCAATTTTGGCTAC	3240
QY	3555	TCCTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCTCCAAAGATTTCAACCT	3614
Db	3241	TCCTGATGACAAACAGGATTTGGACATGTTTCATCAATGCTCCAAACTTCAACCT	3300
QY	3615	CAACATCACCTGGGTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCTCTGT	3674
Db	3301	CAACATCACCTGGGTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCTCTGT	3360
QY	3675	TGTTTCAAAACCAACATTAAGGATGACAAAGATGTTTCTTAATGAGAGTTTGATTT	3734
Db	3361	TGTTTCAAAACCAACATTAAGGATGACAAAGATGTTTCTTAATGAGAGTTTGATTT	3420
QY	3735	TGCGAACCAACCAATATACATTTCTTTGTTTATGTTCAGTAATTTTCACTGCGCCCAACAA	3794
Db	3421	TGCGAACCAACCAATATACATTTCTTTGTTTATGTTCAGTAATTTTCACTGCGCCCAACAA	3480
QY	3795	AATTCAGGT	3803

Db

3481

AATTCAGAT

3489

RESULT 4

EX440935

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX440935

EX440935

CSODP012YL19

5-PRIME, mRNA sequence.

EX440935

EX440935.1

GI:30781889

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10212.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSODP012CF10Q1&cluster=10212.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CSODP012CF10Q1.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODP012YL19"

/tissue type="FETAL BRAIN"

/dev_stage="fetal"

/clone_lib="Homo sapiens FETAL BRAIN"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

FEATURES

source

Query Match

Best Local Similarity

Matches

962;

Conservative

4;

Mismatches

8;

Indels

4;

Gaps

3;

QY

2337

GTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATGCTTGTGTAAGAACTCATG

2396

Db

57

GTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATGCTTGTGTAAGAACTCATG

116

QY

2397

TTTGAAATTTACTACTGCCAAGGAGAAATTATGCAATGCTTAAATTTGTTGTAGGAACCA

2456

Db

117

TTTGAAATTTACTACTGCCAAGGAGAAATTATGCAATGCTTAAATTTGTTGTAGGAACCA

176

QY

2457

CAATGCCCTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTCTTAAGCAGCT

2516

Db

177

CAATGCCCTTTGGCTTCTCTTACACCCCTGATGAAGTGAATTTGCTCTTAAGCAGCT

236

QY

2517

GCGAATATCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGCGCCT

2576

Db

237

GCGAATATCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGCGCCT

296

QY

2577

TCGGAAGATCAATGTGCTTCTGCTGGAGATATGTCCTCCCATTTCAATAGTTT

2636

Db

297

TCGGAAGATCAATGTGCTTCTGCTGGAGATATGTCCTCCCATTTCAATAGTTT

356

QY

2637

ACTACATGGATGCCCTCTGAGCCCATGATGCTGGATTCTGTGGAATTTTATCAGAAC

2696

```

Db 357 ACTACAGTGGATGCGCTCTGAGCCAGTGAAGTCTGTGGAATTTATCAGACCC 416
QY 2697 CAGTACTCGGGAGTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGTGA 2755
Db 417 CAGTACTCGGGAGTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGTGA 476
QY 2757 AAGGCTCTGAAACACAGTCTAAGCAGTCCGGGACACCATGCTGCTTGGAGACAGCATG 2815
Db 477 AAGGCTCTGAAACACAGTCTAAGCAGTCCGGGACACCATGCTGCTTGGAGACAGCATG 536
QY 2817 TGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGT 2876
Db 537 TGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGT 596
QY 2877 GGACTCCAAATGCTATGTGGCTCTCCCTTTGGCCAGTGTATGGAATGTATACAT 2936
Db 597 GGACTCCAAATGCTATGTGGCTCTCCCTTTGGCCAGTGTATGGAATGTATACAT 656
QY 2937 GAGCACTGCCCCCTGAAATTTTCAGGCTACTGTACTGTCTTGTCTTGTGGACA 2995
Db 657 GAGCACTGCCCCCTGAAATTTTCAGGCTACTGTACTGTCTTGTGTGGACA 716
QY 2997 ACCAGCTGTGCTGTGTACTGATCCAGCAATATCTGGCAAGGAAATCATAGAGG 3056
Db 717 ACCAGCTGTGCTGTGTACTGATCCAGCAATATCTGGCAAGGAAATCATAGAGG 776
QY 3057 TTCCTATAAGACACAGTGAAGATGCTTCGACGCCCCCTACAGGAATTTCTATCCACA 3116
Db 777 TTCCTATAAGACACAGTGAAGATGCTTCGACGCCCCCTACAGGAATTTCTATCCACA 836
QY 3117 GCCCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACACTGTCTTTCAITCA 3176
Db 837 GCCCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACACTGTCTTTCAITCA 896
QY 3177 CTGTCCAGTTCCCAATGCAAGGCGCACAGTAATGCAATCAATCAGACATCTGTGAA 3236
Db 897 CTGTCCAGTTCCCAATGCAAGGCGCACAGTAATGCAATCAATCAGACATCTGTGAA 956
QY 3237 GTGTGAGAACCTTGACCAAGGCAAGCACTTGGAGACCTGCAATATCTGGCTTCTACGGTGA 3296
Db 957 GTGTGAGAACCTTGACCAAGGCAAGCACTTGGAGACCTGCAATATCTGGCTTCTACGGTGA 1013
QY 3297 TCCCAACCAATGAGGGAA 3314
Db 1014 T-CCCAACCAATGAGGGAA 1030

```

```

RESULT 5
LOCUS BU506373
DEFINITION AGENCOURT_10015485 NIH_MGC_94 Mus musculus cdna clone IMAGE:6494441
5', mRNA sequence.
ACCESSION BU506373
VERSION BU506373.1 GI:22812606
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 964)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cdna Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

```

Plate: L14M14050 row: k column: 18
 High quality sequence stop: 592.
 Location/Qualifiers
 1. 964
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6494441"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_94"
 Note: "Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

```

Query Match 20.3%; Score 775; DB 13; Length 964;
Best Local Similarity 88.9%; Pred. No. 1.2e-135;
Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;
QY 1996 TGTGTGTGGAAACACAGGCTGTCTCAGTGTATCTCTGGGCGCTGCAACTGATGAACAA 2055
Db 1 TGTCTGTGGACACACAGTGTCTCGATGTACCTCTGGGAGTTGGCAACTGGAACAA 60
QY 2056 GAAGAAAGTTAAATCAGATGTTTTTCCAAAGACCTCTTCACCATCAGATGTGAC 2115
Db 61 GCAGAAAGTTAAATCAGATGTTTTTCTAAAGAACCTTGCACATCAGATGTGAC 120
QY 2116 CAGCACACAGATGTTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTCAATGAC 2175
Db 121 CAGCACACAGATGTTTACAGCTGTACAGCCAAATACCAATGACTGCCACTGGTCAATGAT 180
QY 2176 CATTTGTCTCCCAAGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTGTAGTATGAG 2235
Db 181 CACTGTGTCCCTGTGAAACACACAGCTGTACAGAGGCCAGATCTCCATTTGCCAAGTATGAG 240
QY 2236 AATTGCCCAAGATTAACCCCATGACTACTGTCTTAAAGAACACAGCTGCAGAGAGCTGT 2295
Db 241 AGTTGCCCAAGATTAACCCCATGACTACTGTCTTAAAGAACACAGCTGCAGAGAGCTGT 300
QY 2296 GCCTCGACACAGAACTGCGAGTGGAGCCCGGAATCAGAGTGCATTCCTTGCCTGCGGAA 2355
Db 301 GCCTTAGACAGAACTGCGAGTGGAGCCCGGAATCAGAGTGCATTCCTTGCCTGCGGAA 360
QY 2356 AATATCTGTGGCATTGGCTGGCATTTGTTGAAACTCATGTTGAAATTAATCACTGCTCC 2415
Db 361 AATATCTGTGGCATTGGCTGGCATTTGTTGAAACTCATGTTGAAATTAATCACTGCTCC 420
QY 2416 AAGGAGAAATATGACAAATGCTAAATGTTCTGTAGAACCAACAATGCCCTTTTGTCTTCT 2475
Db 421 AAGGAGAAATATGACAAATGCTAAATGTTCTGTAGAACCAACAATGCCCTTTTGTCTTCC 480
QY 2476 CTTTACCAACCAAGAGAGTGAATTTGTCCTTAGCAGCTGCGAATTAATCAGTCACTCT 2535
Db 481 CTCACATCCCAAGAGAGTGAATTTGTCCTTAGCAGCTTTCGATTAATGCAATCACTCT 540
QY 2536 CAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGCTTCGGGAAGATCAATGTGTCC 2595
Db 541 CAAAGTATGTCCAAGCTCACCTGTACTCTGATCCATGGTGTCTTCGGAAGATCAATGTGTCT 600
QY 2596 TACTGTGTCTGGAGATATGTCCCATTAATAATGTTTACTACAGTGTAGTCCGTCT 2655
Db 601 TACTGTGTCTGGAGATATGTCCCATTAATAATGTTTACTACAGTGTAGTCCGTCT 660
QY 2656 GAGCCAGTGTGTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTC-GGGGACTGAA 2714
Db 661 GAGCCAGTGTGTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTC-GGGGACTGAA 720
QY 2715 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCTGTGAAACCAAG 2774
Db 721 GGCTGCAACCTGCATCAACCCCTCTCAATGGCAGCGCTGTGTGAAAGGCTGTGAAACCAAG 780

```

QY 2775 TGCTAAGCAGTCCGGACACCATGCTGCTTGGAGACAG-CATGTGGAGATTGACACGCG 2833
Db 781 TGCCAAAGCAGTCCGGACACCATGCTGCTTGGAGACAG-CATGTGGAGATTGACACGCG 840
QY 2834 CGAGCTCTGAGTGCATGTGGTGGCAGCAATGAGCAGTGTGAGTCCCAATGCTTATG 2893
Db 841 CCAGCTCGAGTGCATGTGGTGGCAGCAATGAGCAGTGTGAGTCCCAATGCTTATG 900
QY 2894 TGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGATATACGATGAGCAGTGTGAGTCCCAATG 2953
Db 901 TGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGATATACGATGAGCAGTGTGAGTCCCAATG 960
QY 2954 AAAA 2957
Db 961 AAAA 964

RESULT 6
Bg678679
LOCUS 602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
DEFINITION mRNA sequence.
ACCESSION Bg678679
VERSION Bg678679.1 GI:13910076
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10602 row: j column: 20
High quality sequence stop: 762.
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="PH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI CGAP Library."

FEATURES
source
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="PH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 20.0%; Score 763.2; DB 12; Length 933;
Best Local Similarity 97.3%; Pred. No. 1.9e-133;
Matches 798; Conservative 0; Mismatches 18; Indels 4; Gaps 2;

QY 2997 ACCAGGCTGGCTGGTGTACTGATCCAGCAATCTGCGCAAGGAAATGATAGAGGG 3056
Db 2 ACCAGGCTGGCTGGTGTACTGATCCAGCAATCTGCGCAAGGAAATGATAGAGGG 61
QY 3057 TTCCCTATAAGGACCAAGTGAAGTGCCTTGGCAAGCCCTACAGGAAATTTCTATCCCA 3116
Db 62 TTCCCTATAAGGACCAAGTGAAGTGCCTTGGCAAGCCCTACAGGAAATTTCTATCCCA 121
QY 3117 GCCCTGCTCAATTCAGCATGTGTAGAGGACAGCAGATACCACTGCTTTCATTCA 3176

Db 122 GCCCTGCTCAATTCAGCATGTGTAGAGGACAGCAGATCAACTGCTTTCATTCA 181
QY 3177 CTGTCAGGCTTGCATATGCAACGGCCACAGTAAATGATCAATCAAGATCTGTGAGAA 3236
Db 182 CTGTCAGGCTTGCATATGCAACGGCCACAGTAAATGATCAATCAAGATCTGTGAGAA 241
QY 3237 GTGTGAGAACCTGACACAGGAGCACTGCGAGACCTCATATCTGGCTTCTACGGTGA 3296
Db 242 GTGTGAGAACCTGACACAGGAGCACTGCGAGACCTCATATCTGGCTTCTACGGTGA 301
QY 3297 TCCCAACAATGAGGAGAAATGTCAGCCATGCAAGTGCATATGCGCAGCGCTCTGTGCAA 3356
Db 302 TCCCAACAATGAGGAGAAATGTCAGCCATGCAAGTGCATATGCGCAGCGCTCTGTGCAA 361
QY 3357 CACCAACAGGAGCAAGTGTCTGTCACCAAGGAGCGCTCAAGGAGGAGGAGGCCAGCT 3416
Db 362 CACCAACAGGAGCAAGTGTCTGTCACCAAGGAGCGCTCAAGGAGGAGGAGGCCAGCT 421
QY 3417 ATGTGAGGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATATCTCT 3476
Db 422 ATGTGAGGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATATCTCT 481
QY 3477 TCTTATGACTATCAGTTACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGC 3536
Db 482 TCTTATGACTATCAGTTACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGC 541
QY 3537 TATCAATTTTGTGCTACTCTCTGACGAAACAAACAGGAGTTTCGACATGTTCAATATGC 3596
Db 542 TATCAATTTTGTGCTACTCTCTGACGAAACAAACAGGAGTTTCGACATGTTCAATATGC 601
QY 3597 CTCCAAGAAATTTCAACCTCAACATCACTGGGCTGCGAGTTTCTAGCTGGAACCCAGGC 3656
Db 602 CTCCAAGAAATTTCAACCTCAACATCACTGGGCTGCGAGTTTCTAGCTGGAACCCAGGC 661
QY 3657 TGGAGAGAGATGCTGTTGTTTTCACAAACCAACATAGGACTACAAAGATAGTTTCTC 3716
Db 662 TGGAGAGAGATGCTGTTGTTTTCACAAACCAACATAGGACTACAAAGATAGTTTCTC 721
QY 3717 TAATGAGAGTTTCAATTTTCGCAACCAACCAATATACCTTTCTTTTATGTCAGTAA 3776
Db 722 TAATGAGAGTTTCAATTTTCGCAACCAACCAATATACCTTTCTTTTATGTCAGTAA 780
QY 3777 TTTCACCT---GGCCCATCAAAATTCAGGTGCAAACTGAA 3813
Db 781 TTGTCACCTTGGGCCCCATCGAAATTCAGGTGCAAAATGGA 820

RESULT 7
CB520374
LOCUS 840 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GIO-cei-1-18-0-UT.r1 NIH BMAP_GIO Mus musculus cDNA clone
IMAGE:6840067 5', mRNA sequence.
ACCESSION CB520374
VERSION CB520374.1 GI:29353729
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousef1.html
This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1. 840
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840067"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
/lab_host="PH108 (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pyx-Asc; Site 1: EcoR I; Site 2: NotI; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN
Query Match 18.4%; Score 701.2; DB 14; Length 840;
Best Local Similarity 90.2%; Pred. No. 9.5e-122;
Matches 759; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 2151 CAATGACTGCCACTGTGCAATGACATTTGTCTCCAGGACACACACTGCTCAGAGG 2210
DB 1 CAATGACTGCCACTGTGCAATGACATTTGTCTCCAGGACACACACTGCTCAGAGG 60

QY 2211 CCAGATCTCCATTTTAGTGTATGAAATTTGCCCAAGGATACCCCATGCTACTGTAA 2270
DB 61 CCAGATCTCCATTTTAGTGTATGAAATTTGCCCAAGGATACCCCATGCTACTGTAA 120

QY 2271 CAAGAAGACCACTGCAGAGCTGTGCCCTGGACAGCAACTGCCAGTGGGACCCCGAA 2330
DB 121 TAAGAAGACCACTGCAGAGCTGTGCCCTGGACAGCAACTGCCAGTGGGACCCCGAA 180

QY 2331 TCAGGAGTGCATTCGCTCCCGCAAAATATCTGTGGCAATTCGTCGCAATTTGTTGAA 2390
DB 181 TCAGGAGTGCATTCGCTCCCGCAAAATATCTGTGGCAATTCGTCGCAATTTGTTGAA 240

QY 2391 CTCATGTTTGAATTAATCTACTGCCAAGGAAATATGACAAATGCTAAATTTGTTCTAG 2450
DB 241 CTCGTTCTGAAATCACTACTGCTAAGGAAATATGACAAATGCTAAATTTGTTCTAG 300

QY 2451 GAACACAAATGCCCTTTTGGCTTCTTTAACCAGCAAGAGTAGAATTTGTCCTTAA 2510
DB 301 GAACACAAATGCCCTTTTGGCTTCTTTAACCAGCAAGAGTAGAATTTGTCCTTAA 360

QY 2511 GCAGCTGCGAATAATGCAAGTATCTCAGAGCATGCCAAGTCACTTTAACCCCATGGT 2570
DB 361 GCAGCTGCGAATAATGCAAGTATCTCAGAGCATGCCAAGTCACTTTAACCCCATGGT 420

QY 2571 CGGCTTCGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGCCCATTAACAA 2630
DB 421 TGGTCTTCGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGCCCATTAACAA 480

QY 2631 TAGTTTACTACAGTGCATGCGCTGTGAGCCAGTCACTGATTCGTGCAATTTATC 2690
DB 481 TAGTTTACTACAGTGCATGCGCTGTGAGCCAGTCACTGATTCGTGCAATTTATC 540

QY 2691 AGAACCCAGTACTCGGAGCTGAAGGCTGCAACCTGTCATCAACCCACTCAATGTTAGT 2750
DB 541 AGAACCCAGTACTCGGAGCTGAAGGCTGCAACCTGTCATCAACCCACTCAATGTTAGT 600

QY 2751 CTGTGAAGGCGCTGCAAAACACACAGTCTAAGCAGTGGCCGACACCATGTGCTTGGAGC 2810
DB 601 CTGTGAAGGCGCTGCAAAACACACAGTCTAAGCAGTGGCCGACACCATGTGCTTGGAGC 660

QY 2811 AGCATGTGGAGATTGACACAGCGGCGAGCTCTGAGTGCATGTGTGACAGCAACATGGTA 2870
DB 661 AGCGTGTGGGAGTGCACCTAGCAGCAGCTCGAGTGCATGTGTGAGTACATGAAGCA 720

QY 2871 GTGTGTGCACTCAATGCCCTATGGCTCTCTCCCTTTTGGCCAGTGTATGAATGGTA 2930
DB 721 GTGTGTGCACTCAATGCCCTATGGCTCTCTCCCTTTTGGCCAGTGTATGAATGGTA 780

QY 2931 TACGATCAGACCTGCCCCCTGAAAATTTTCAGGCTACTGTACCTGTAGTCAATGCTT 2990
DB 781 TACGATCAGACCTGCCCCCTGAAAATTTTCAGGCTACTGTACCTGTAGTCAATGCTT 839

QY 2991 G 2991
DB 840 G 840

RESULT 8
BM783739
LOCUS K-EST0061730 S5SNU484 Homo sapiens CDNA clone S5SNU484-34-H05 5', linear EST 05-MAR-2002
DEFINITION mRNA sequence.
ACCESSION BM783739
VERSION BM783739
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yusong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: H column: 05
High quality sequence stop: 681.

FEATURES
source
1. 681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-H05"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P by electroporation

ORIGIN		method."	
Query Match		17.8%;	Score 681; DB 12; Length 681;
Best Local Similarity		100.0%;	Pred. No. 5.8e-118;
Matches 681;		Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	2512	CAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGCTCAAGCTCAACCTTAAACCCCATGGGTC	2571
DB	1	CAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGCTCAAGCTCAACCTTAAACCCCATGGGTC	60
QY	2572	GGCCTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGCCCATTTTACAAAT	2631
DB	61	GGCCTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGCCCATTTTACAAAT	120
QY	2632	AGTTTACTACAGTGGATGCGCTCTGAGCCAGTCAATGCTGGATCTGTGGAAATTTTATCA	2691
DB	121	AGTTTACTACAGTGGATGCGCTCTGAGCCAGTCAATGCTGGATCTGTGGAAATTTTATCA	180
QY	2692	GAACCCAGTACTCGGAGCTGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGC	2751
DB	181	GAACCCAGTACTCGGAGCTGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGC	240
QY	2752	TGTCGAAGGCTGCAACCAACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGCA	2811
DB	241	TGTCGAAGGCTGCAACCAACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGCA	300
QY	2812	GCATGTGAGATTCGACAGCGGAGCTCTGAGTGCATGTGGTGCGACCAATGAAGCAG	2871
DB	301	GCATGTGAGATTCGACAGCGGAGCTCTGAGTGCATGTGGTGCGACCAATGAAGCAG	360
QY	2872	TGCTGGACTCCCAATGCTATGCTGCTCTTCCCTTTGGCCAGTGTATGGAAATGGTAT	2931
DB	361	TGCTGGACTCCCAATGCTATGCTGCTCTTCCCTTTGGCCAGTGTATGGAAATGGTAT	420
QY	2932	ACGATGAGCACTGCCCCCTGCAAAATGTTTCAGGCTACTGTACCTCTAGTCAATGGTTC	2991
DB	421	ACGATGAGCACTGCCCCCTGCAAAATGTTTCAGGCTACTGTACCTCTAGTCAATGGTTC	480
ORIGIN		Query Match	
Best Local Similarity		17.8%;	Score 680.4; DB 12; Length 683;
Matches 681;		Conservative 0;	Mismatches 2; Indels 0; Gaps 0;
QY	2512	CAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGCTCAAGCTCAACCTTAAACCCCATGGGTC	2571
DB	1	CAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGCTCAAGCTCAACCTTAAACCCCATGGGTC	60
QY	2572	GGCCTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGCCCATTTTACAAAT	2631
DB	61	GGCCTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGCCCATTTTACAAAT	120
QY	2632	AGTTTACTACAGTGGATGCGCTCTGAGCCAGTCAATGCTGGATCTGTGGAAATTTTATCA	2691
DB	121	AGTTTACTACAGTGGATGCGCTCTGAGCCAGTCAATGCTGGATCTGTGGAAATTTTATCA	180
QY	2692	GAACCCAGTACTCGGAGCTGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGC	2751
DB	181	GAACCCAGTACTCGGAGCTGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGC	240
QY	2752	TGTCGAAGGCTGCAACCAACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGCA	2811
DB	241	TGTCGAAGGCTGCAACCAACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGCA	300
QY	2812	GCATGTGAGATTCGACAGCGGAGCTCTGAGTGCATGTGGTGCGACCAATGAAGCAG	2871
DB	301	GCATGTGAGATTCGACAGCGGAGCTCTGAGTGCATGTGGTGCGACCAATGAAGCAG	360
QY	2872	TGCTGGACTCCCAATGCTATGCTGCTCTTCCCTTTGGCCAGTGTATGGAAATGGTAT	2931
DB	361	TGCTGGACTCCCAATGCTATGCTGCTCTTCCCTTTGGCCAGTGTATGGAAATGGTAT	420
QY	2932	ACGATGAGCACTGCCCCCTGCAAAATGTTTCAGGCTACTGTACCTCTAGTCAATGGTTC	2991
DB	421	ACGATGAGCACTGCCCCCTGCAAAATGTTTCAGGCTACTGTACCTCTAGTCAATGGTTC	480

TITLE 21C Frontier Korean EST Project 2001

```

QY 2992 GAGCAACCGGCTGTGGTGTACTGATCCAGCAATATCGGCAAGGGAATGCATA 3051
    |||||
Db 481 GAGCAACCGGCTGTGGTGTACTGATCCAGCAATATCGGCAAGGGAATGCATA 540
    |||||
QY 3052 GAGGGTTCCTATAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 3111
    |||||
Db 541 GAGGGTTCCTATAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 600
    |||||
QY 3112 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTC 3171
    |||||
Db 601 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTC 660
    |||||
QY 3172 ATTCACTGCCAGCTTGCCAATG 3194
    |||||
Db 661 ATTCACTGCCAGCTTGCCAATG 683
    |||||

RESULT 10
BM783258
LOCUS K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
DEFINITION mRNA sequence.
ACCESSION BM783258
VERSION BM783258.1 GI:19131490
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 31 row: H column: 08
High quality sequence stop: 680.
Location/Qualifiers
1. 680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_lines="SNU-484"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo di-selected
mRNA by priming with dr-tailed vector. The dr-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P' by electroporation
method."

FEATURES
source
1. 680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_lines="SNU-484"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo di-selected
mRNA by priming with dr-tailed vector. The dr-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P' by electroporation
method."

ORIGIN
Query Match 17.8%; Score 680; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 8.9e-118;

```

```

Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2512 CAGCTCGGAATAATCAGATCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTTC 2571
    |||||
Db 1 CAGCTCGGAATAATCAGATCATCTCAGAGCATGTCCTCAAGCTCACCTTAACCCCATGGTTC 60
    |||||
QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTTACAAAT 2631
    |||||
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTTACAAAT 120
    |||||
QY 2632 AGTTTACTACAGTGGATCGCGTCTGAGCCCACTGATGCTGGATCTCTGCGAATTTATCA 2691
    |||||
Db 121 AGTTTACTACAGTGGATCGCGTCTGAGCCCACTGATGCTGGATCTCTGCGAATTTATCA 180
    |||||
QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGCTAGTGTTC 2751
    |||||
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGCTAGTGTTC 240
    |||||
QY 2752 TGTGAAGGCTGCAACCCACACAGTGTCTAAGCAGTCCCGGACACCATGTGCTTGGAGCA 2811
    |||||
Db 241 TGTGAAGGCTGCAACCCACACAGTGTCTAAGCAGTCCCGGACACCATGTGCTTGGAGCA 300
    |||||
QY 2812 GCATGTGGAGATTGCACACGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG 2871
    |||||
Db 301 GCATGTGGAGATTGCACACGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG 360
    |||||
QY 2872 TGTGTGCACTCCAATGCCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2931
    |||||
Db 361 TGTGTGCACTCCAATGCCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
    |||||
QY 2932 ACATGACGACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTAGTCTGTG 2991
    |||||
Db 421 ACATGACGACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTAGTCTGTG 480
    |||||
QY 2992 GAGCAACCGGCTGTGGTGTACTGATCCAGCAATATCGGCAAGGGAATGCATA 3051
    |||||
Db 481 GAGCAACCGGCTGTGGTGTACTGATCCAGCAATATCGGCAAGGGAATGCATA 540
    |||||
QY 3052 GAGGGTTCCTATAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 3111
    |||||
Db 541 GAGGGTTCCTATAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 600
    |||||
QY 3112 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTC 3171
    |||||
Db 601 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTC 660
    |||||
QY 3172 ATTCACTGCCAGCTTGCCA 3191
    |||||
Db 661 ATTCACTGCCAGCTTGCCA 680
    |||||

```

```

RESULT 11
BM783773
LOCUS K-EST00611775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
DEFINITION tRNA sequence.
ACCESSION BM783773
VERSION BM783773.1 GI:19132005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

```

Tel: +82-42-860-4470
 Fax: +82-42-860-4470
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: D column: 02
 High quality sequence stop: 680.
 Location/Qualifiers
 1. 680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-36-D02"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /lab_host="SNU-484"
 /lab_host="Top10F"
 /clone_lib="S5SNU484"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-tailed mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

FEATURES
 source
 1. 680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-34-F01"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /lab_host="Top10F"
 /clone_lib="S5SNU484"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

ORIGIN
 Query Match 17.8%; Score 680; DB 12; Length 680;
 Best Local Similarity 100.0%; Pred. No. 8.9e-118;
 Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACGCTCAGCTTAAACCCCATGGGTC 2571
 DB 1 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACGCTCAGCTTAAACCCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTGTCTTACTGTGTGCTGGGAAGATATGTCCTCCCATTTACAAAT 2631
 DB 61 GGCCTTCGGAAGATCAATGTGTCTTACTGTGTGCTGGGAAGATATGTCCTCCCATTTACAAAT 120

QY 2632 AGTTTACTACAGTGGATCCCTCTGAGCCAGTGTGATCTGTGGATTTTATCA 2691
 DB 121 AGTTTACTACAGTGGATCCCTCTGAGCCAGTGTGATCTGTGGATTTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCT 2751
 DB 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCT 240

QY 2752 TGTGAAGGCTGCAACACAGTGTCTAAGCAGTCCCGACACCATGTCCTTGGAGCA 2811
 DB 241 TGTGAAGGCTGCAACACAGTGTCTAAGCAGTCCCGACACCATGTCCTTGGAGCA 300

QY 2812 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGAGCAACATCAAGCAG 2871
 DB 301 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGAGCAACATCAAGCAG 360

QY 2872 TGTGTGAGTCCAAATGCGCTATGTGCGCTCCCTTCCCTTTGGCCAGTGTATGGATGGTAT 2931
 DB 361 TGTGTGAGTCCAAATGCGCTATGTGCGCTCCCTTCCCTTTGGCCAGTGTATGGATGGTAT 420

QY 2932 ACGATGAGCACTGCCCCCTCGAAATGTTTCAGGCTACTGTACCTGTAGTCATGCTTG 2991
 DB 421 ACGATGAGCACTGCCCCCTCGAAATGTTTCAGGCTACTGTACCTGTAGTCATGCTTG 480

QY 2992 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATCTGGCAAAAGGAAATGCATA 3051
 DB 481 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATCTGGCAAAAGGAAATGCATA 540

QY 3052 GAGGCTTCTATAAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTTACAGGAAATTTCTAT 3111

RESULTS
 BM783718 677 bp mRNA linear EST 05-MAR-2002
 LOCUS K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',
 DEFINITION mRNA sequence.
 ACCESSION BM783718
 VERSION BM783718.1 GI:19131950
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 677)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Kim N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 CONTACT: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: F column: 01
 High quality sequence stop: 677.
 Location/Qualifiers
 1. 677
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-34-F01"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /lab_host="Top10F"
 /clone_lib="S5SNU484"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

ORIGIN
 Query Match 17.7%; Score 677; DB 12; Length 677;
 Best Local Similarity 100.0%; Pred. No. 3.3e-117;
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACGCTCAGCTTAAACCCCATGGGTC 2571
 DB 1 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACGCTCAGCTTAAACCCCATGGGTC 60

FEATURES		Location/Qualifiers	
source		1. 677	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="S5SNU484-36-H09"	
		/sex="M"	
		/tissue_type="Stomach"	
		/cell_type="Epithelial"	
		/cell_line="SNU-484"	
		/lab_host="Top10F"	
		/clone_lib="S5SNU484"	
		/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method."	
ORIGIN		Query Match 17.7%; Score 674.4; DB 12; Length 677;	
		Best Local Similarity 99.7%; Pred. No. 1e-116;	
		Matches 675; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	2512	CAGCTCGGAATTAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC	2571
Db	1	CAGCTCGGAATTAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC	60
QY	2572	GGCTTTGGAAGATCAATGTGCTTCTGCTGGGAAGATATGTCCTCAATTTACAAAT	2631
Db	61	GGCTTTGGAAGATCAATGTGCTTCTGCTGGGAAGATATGTCCTCAATTTACAAAT	120
QY	2632	AGTTTACTACGTGATGCGCTGTGAGCCAGTGTGTTGATTTCTGTTGGAATTTATCA	2691
Db	121	AGTTTACTACGTGATGCGCTGTGAGCCAGTGTGTTGATTTCTGTTGGAATTTATCA	180
QY	2692	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	2751
Db	181	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240
QY	2752	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCTTGGAGCA	2811
Db	241	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCTTGGAGCA	300
QY	2812	GCATGTGGAGATTGCAACCCAGCGGACGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG	2871
Db	301	GCATGTGGAGATTGCAACCCAGCGGACGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG	360
QY	2872	TGTGTGACTCAATGCCTATGTGGCTTCTTCCCTTTTGGCCAGTGTATGGAAATGGTAT	2931
Db	361	TGTGTGACTCAATGCCTATGTGGCTTCTTCCCTTTTGGCCAGTGTATGGAAATGGTAT	420
QY	2932	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTTGTG	2991
Db	421	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTTGTG	480
QY	2992	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATGCGCAAGGAAATGCATA	3051
Db	481	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATGCGCAAGGAAATGCATA	540
QY	3052	GAGGTTCTTATAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTAT	3111
Db	541	GAGGTTCTTATAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTAT	600
QY	3112	CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGGACAGCAGATCAACCTGGTCTTTC	3171

QY	2572	GGCTTCGGAAGATCAATGTGCTTCTACTGCTGGGAAGATATGTCCTCCATTTTACAAT	2631
Db	61	GGCTTCGGAAGATCAATGTGCTTCTACTGCTGGGAAGATATGTCCTCCATTTTACAAT	120
QY	2632	AGTTTACTACGTGATGCGCTGTGAGCCAGTGTGTTGATTTCTGTTGGAATTTATCA	2691
Db	121	AGTTTACTACGTGATGCGCTGTGAGCCAGTGTGTTGATTTCTGTTGGAATTTATCA	180
QY	2692	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	2751
Db	181	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240
QY	2752	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCTTGGAGCA	2811
Db	241	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCTTGGAGCA	300
QY	2812	GCATGTGGAGATTGCAACCCAGCGGACGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG	2871
Db	301	GCATGTGGAGATTGCAACCCAGCGGACGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG	360
QY	2872	TGTGTGACTCAATGCCTATGTGGCTTCTTCCCTTTTGGCCAGTGTATGGAAATGGTAT	2931
Db	361	TGTGTGACTCAATGCCTATGTGGCTTCTTCCCTTTTGGCCAGTGTATGGAAATGGTAT	420
QY	2932	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTTGTG	2991
Db	421	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTTGTG	480
QY	2992	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATGCGCAAGGAAATGCATA	3051
Db	481	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATGCGCAAGGAAATGCATA	540
QY	3052	GAGGTTCTTATAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTAT	3111
Db	541	GAGGTTCTTATAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTAT	600
QY	3112	CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGGACAGCAGATCAACCTGGTCTTTC	3171
Db	601	CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGGACAGCAGATCAACCTGGTCTTTC	660
QY	3172	ATTCACTGTCCAGCTTG 3188	
Db	661	ATTCACTGTCCAGCTTG 677	

RESULT 13
 BM783821
 LOCUS K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5',
 DEFINITION mRNA sequence.
 ACCESSION BM783821
 VERSION BM783821.1 GI:19132053
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 677)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Kim,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: H column: 09
 High quality sequence stop: 677.

```
Db 601 CCACAGCCCTGCTCAATCCAGATGCTCTAGAGGACAGACAGATACACTGCTCTTTC 660
QY 3172 ATTACTGTCAGCTTG 3188
Db 661 ATTACTGTCAGCTTG 677

RESULT 14
BM783685
LOCUS K-EST0061659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5',
DEFINITION mRNA sequence.
ACCESSION BM783685
VERSION EST.
KEYWORDS BM783685.1 GI:19131917
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-gong yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: B column: 06
High quality sequence stop: 673.
FEATURES
location/Qualifiers
1..673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-B06"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18Rpl; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F, by electroporation
method."

ORIGIN
Query Match 17.6%; Score 673; DB 12; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.9e-116; Mismatches 0; Indels 0; Gaps 0;
Matches 673; Conservative 0;

QY 2512 CAGCTCGGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 2571
Db 1 CAGCTCGGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCATGTGTCCTACTGCTGGAGAGATATGTCCCATTTACAAT 2631
Db 61 GGCCTTCGGAAGATCATGTGTCCTACTGCTGGAGAGATATGTCCCATTTACAAT 120
```

```
QY 2632 AGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGAAGTCTGGATTCTGTGGAATTTATCA 2691
Db 121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGAAGTCTGTGGAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTTC 2751
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTTC 240

QY 2752 TGTGAAGGGCTCTCAACACACAGTGTAAAGCAGTGCAGGACACCATGTGCTTGGAGACA 2811
Db 241 TGTGAAGGGCTCTCAACACACAGTGTAAAGCAGTGCAGGACACCATGTGCTTGGAGACA 300

QY 2812 GCATGTGGAGATTGCACACGCGGACGCTCTGAGTGCATGTGTTGAGACACATGAAGCAG 2871
Db 301 GCATGTGGAGATTGCACACGCGGACGCTCTGAGTGCATGTGTTGAGACACATGAAGCAG 360

QY 2872 TGTGTGACTCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGACTCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

QY 2932 ACGATGAGCAGCTGCCCCCTGAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACGATGAGCAGCTGCCCCCTGAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480

QY 2992 GAGCAACGAGCTGTGGCTGCTGCTACTGATCCAGCAATACTGCGAAGGGAATGCATA 3051
Db 481 GAGCAACGAGCTGTGGCTGCTGCTACTGATCCAGCAATACTGCGAAGGGAATGCATA 540

QY 3052 GAGGTTTCCTTATAAGGACCAAGATGCTTCCCAAGCCCTACAGGAAATTTCTAT 3111
Db 541 GAGGTTTCCTTATAAGGACCAAGATGCTTCCCAAGCCCTACAGGAAATTTCTAT 600

QY 3112 CCACAGCCCTGCTCAATTCAGATGTGTCTAGAGACAGCAGATACACTGGTCTTTC 3171
Db 601 CCACAGCCCTGCTCAATTCAGATGTGTCTAGAGACAGCAGATACACTGGTCTTTC 660

QY 3172 ATTCACTGTCCAG 3184
Db 661 ATTCACTGTCCAG 673

RESULT 15
BX431593
LOCUS BX431593
DEFINITION BX431593 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSODF012YL19 5-PRIME, mRNA sequence.
ACCESSION BX431593
VERSION BX431593.1 GI:30781048
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 962)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10212.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG017ZB02_CS01546_1&cluster=10212.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG017ZB02_CS01546_1.
FEATURES
location/Qualifiers
1..962
/organism="Homo sapiens"
/mol_type="mRNA"
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:06:34 ; Search time 899.913 Seconds
(without alignments)
18028.269 Million cell updates/sec

Title: US-09-787-097-19
Perfect score: 3819
Sequence: 1 atggtgcccagcgccgccc.....aggtgcaactgaacaatga 3819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_29Jan04.*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002s.*
 - 7: Geneseqn2003as.*
 - 8: Geneseqn2003bs.*
 - 9: Geneseqn2003cs.*
 - 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3819	100.0	3819	3	Aaz52277 Human sol
2	3801.4	99.5	4290	3	Aaz52276 Human mem
3	3801.4	99.5	8561	5	Aaz522659 DNA encod
4	3713.8	97.2	4072	3	Aaz91921 Human mah
5	3696.2	96.8	8590	3	Aaz91920 Human mah
6	3492.4	91.4	4068	3	Aaz52275 Human mem
7	3487	91.3	3597	3	Aaz52274 Human sol
8	3487	91.3	3597	5	Aaz522657 DNA encod
9	3487	91.3	3597	9	ADb70997 Human att
10	3114.8	81.6	8739	9	ADb53249 Primary r
11	2338.2	76.9	8827	3	Aaz91916 Wild type
12	2210.2	73.6	3490	5	Aaz522660 DNA encod
13	2248.2	58.9	2625	3	Aaz91922 Human mah
14	1521.8	39.8	2419	3	Aaz91917 Murine mah
15	1341.6	35.1	6373	3	Aaz91919 Human mah
16	1233.6	32.3	4134	7	ABz58710 Murine ho
17	1228.2	32.2	4140	7	ABz58709 Human ho
18	1135	29.7	3609	6	ABs63434 Human cDN
19	873.2	22.9	3165	5	Aaz58781 DNA encod
20	873.2	22.9	3313	5	Aaz58782 DNA encod
21	819.4	21.5	1012	2	Aax84568 Biorhythm
22	640.2	16.8	1051	3	Aaz91918 Murine mah
23	574	15.0	6517	4	Aah98748 Human EST

24	536	14.0	6532	5	Aas86766 DNA encod
25	421.8	11.0	671	5	Aas72658 DNA encod
26	420	11.0	1654	7	ABx34730 Human mad
27	415.2	10.9	1383	6	AAL54064 CDNA of h
28	332	8.7	668	4	AAS02409 Human sec
29	277.4	7.3	481	4	AAL14686 Human bre
30	276.4	7.2	292	3	AAA45505 Human sec
31	267.6	7.0	845	6	ABQ52394 Oligonuc1
32	267.6	7.0	845	6	ABQ52395 Oligonuc1
33	266.4	7.0	482	4	AAL23548 Human bre
34	253.8	6.6	982	6	ABL39784 Human NS
35	249.2	6.5	531	4	AAL143425 Probe #12
36	249.2	6.5	531	6	ABs11523 Human gen
37	246	6.4	246	6	ABK38928 cDNA enco
38	246	6.4	246	7	ACal11257 Human lun
39	246	6.4	246	7	ACA02443 Lung canc
40	236	6.2	657	7	ABz58714 Upregulat
41	234.4	6.1	207433	5	ABz72040 Gene 216
42	234.4	6.1	207433	7	ABx74891 BAC1098L2
43	209.8	5.5	5312	6	ABL32196 Human imm
44	206	5.4	845	6	ABQ52396 Oligonuc1
45	206	5.4	845	6	ABQ52397 Oligonuc1

ALIGNMENTS

RESULT 1
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.

XX AC AAZ52277;
XX AC

DT 18-JUL-2000 (first entry)
XX Human soluble attractin-2 cDNA.

XX Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 1..3819
FT /*tag= a
FT /product= "Soluble attractin-2"

XX WO200015651-A1.
XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-US020948.
XX 14-SEP-1998; 98US-0100137P.

XX (DAND) DANA FARBER CANCER INST INC.
XX Duke-Cohan JS, Schlossman SF;
XX WPI; 2000-271373/23.

XX P-PSDB; AAY70692.
XX Isolated nucleic acids encoding human attractin polypeptides useful for
XX enhancing immune responses.

XX Claim 3; Fig 12; 120pp; English.
XX The patent discloses four forms of human attractin polypeptides which
XX enhance immune response by promoting macrophage and monocyte spreading in
XX the presence of T cells. These include soluble attractin-1 and -2 and
XX membrane attractin-1 and -2. These various forms of attractin are encoded
XX by alternatively spliced mRNA molecule transcribed from a single gene.

CC The present sequence is a cDNA encoding human soluble attractin-2.
 CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
 CC the N-terminal. This sequence can be used to enhance immune response in
 CC immunosuppressed patients such as those undergoing chemo- and radio-
 CC therapy treatment for cancer or those suffering from common variable
 CC immunodeficiency syndrome. The proteins may also be used to screen
 CC modulators (agonists and antagonists) of immune response which may also
 CC be used to regulate immune reactions. Attractin antibodies can be used to
 CC inhibit immune response in transplant recipients or patients afflicted
 CC with autoimmune disease
 XX
 SQ

Sequence 3819 Bp; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

Query Match 100.0%; Score 3819; DB 3; Length 3819;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3819; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGGCGCAGCGCGCACTGAGCAAGGCTGAGGAGGAGGACGGCGCGACGCA 60
 DB |||||
 QY 1 ATGGTGGCGCAGCGCGCACTGAGCAAGGCTGAGGAGGAGGACGGCGCGACGCA 60
 DB |||||
 QY 61 GCGCTCGCGGCGAGGACGGCGCGCACTGGGACGTGCGACGTGCGAGGCTGGGAGG 120
 DB |||||
 QY 61 GCGCTCGCGGCGAGGACGGCGCGCACTGGGACGTGCGACGTGCGAGGCTGGGAGG 120
 DB |||||
 QY 121 CCGGGCTGGGGCGGGCTGGCGCTCCCGCGGCTGCTGCTCCACCGCTGGGCGACG 180
 DB |||||
 QY 121 CCGGGCTGGGGCGGGCTGGCGCTCCCGCGGCTGCTGCTCCACCGCTGGGCGACG 180
 DB |||||
 QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB |||||
 QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB |||||
 QY 241 GCGAGGCGCGCGCGCGCGCGCGGCTGCGGCTCAGCGCGACCGCGCGCAAGGAA 300
 DB |||||
 QY 241 GCGAGGCGCGCGCGCGCGCGCGGCTGCGGCTCAGCGCGACCGCGCGCAAGGAA 300
 DB |||||
 QY 301 TGTGACCGCGCTGCTGCTCAACCGCGCTGCTGCTCAACCGCGCTGCTGCTGCTGCTG 360
 DB |||||
 QY 301 TGTGACCGCGCTGCTGCTCAACCGCGCTGCTGCTCAACCGCGCTGCTGCTGCTGCTG 360
 DB |||||
 QY 361 CCGCGCGGCTGGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAACTGGA 420
 DB |||||
 QY 361 CCGCGCGGCTGGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAACTGGA 420
 DB |||||
 QY 421 TCTTCTGGGTTTGTGACAGATGCACTGGAAATTAATAATACAAACGAAGTGCACGTGG 480
 DB |||||
 QY 421 TCTTCTGGGTTTGTGACAGATGCACTGGAAATTAATAATACAAACGAAGTGCACGTGG 480
 DB |||||
 QY 481 CTCATTGAAGGACAGCCAAATAGAAATAGACTTCGTTTCAATCATTTTGTCTACAGAG 540
 DB |||||
 QY 481 CTCATTGAAGGACAGCCAAATAGAAATAGACTTCGTTTCAATCATTTTGTCTACAGAG 540
 DB |||||
 QY 541 TGTAGTTGGACCAATTTATATGTTTATGAGGGGACTCAATTTATGCAACCGCTAGTTGCT 600
 DB |||||
 QY 541 TGTAGTTGGACCAATTTATATGTTTATGAGGGGACTCAATTTATGCAACCGCTAGTTGCT 600
 DB |||||
 QY 601 GCATTTAGTGGCTCATTTGTTTCTCAGAGAGATGCAATGAGACTGTCCTGAGGTTGTT 660
 DB |||||
 QY 601 GCATTTAGTGGCTCATTTGTTTCTCAGAGAGATGCAATGAGACTGTCCTGAGGTTGTT 660
 DB |||||
 QY 661 GCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 DB |||||
 QY 661 GCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 DB |||||
 QY 721 TTTTATATTTACTTACAGTTTGTATGTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 DB |||||
 QY 721 TTTTATATTTACTTACAGTTTGTATGTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 DB |||||
 QY 781 AAGATCAGTAATAGCAGCGAACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTTGAA 840
 DB |||||
 QY 781 AAGATCAGTAATAGCAGCGAACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTTGAA 840
 DB |||||

QY 841 GCATGTGACATTTCTCTACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCACTTGCAT 900
 DB |||||
 QY 841 GCATGTGACATTTCTCTACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCACTTGCAT 900
 DB |||||
 QY 901 TCAAGTGTAGTGTACAGGATGCTGCTGCTCTCAGACTGGCAGGGTCTGTGATTTTCAGTT 960
 DB |||||
 QY 901 TCAAGTGTAGTGTACAGGATGCTGCTGCTCTCAGACTGGCAGGGTCTGTGATTTTCAGTT 960
 DB |||||
 QY 961 CCTGTACACGTAACCACTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCC 1020
 DB |||||
 QY 961 CCTGTACACGTAACCACTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCC 1020
 DB |||||
 QY 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACAATTTGTGGGTTGTGTGAGGATATATG 1080
 DB |||||
 QY 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACAATTTGTGGGTTGTGTGAGGATATATG 1080
 DB |||||
 QY 1081 TTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTT 1140
 DB |||||
 QY 1081 TTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTT 1140
 DB |||||
 QY 1141 CCACCTAACCGTTCTCTGTGAACAATGTGGTTGTAGATATGTTCTTTTGGCAATTATAC 1200
 DB |||||
 QY 1141 CCACCTAACCGTTCTCTGTGAACAATGTGGTTGTAGATATGTTCTTTTGGCAATTATAC 1200
 DB |||||
 QY 1201 AAGGATTAATAATTTACATGTATGGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260
 DB |||||
 QY 1201 AAGGATTAATAATTTACATGTATGGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260
 DB |||||
 QY 1261 TTGAGAGTTTTCACATTAATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1320
 DB |||||
 QY 1261 TTGAGAGTTTTCACATTAATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1320
 DB |||||
 QY 1321 CAGTATGCACTGTTGGGCACTCTGCAACATTTTACACTGAAGAATGGCCAGTGGTC 1380
 DB |||||
 QY 1321 CAGTATGCACTGTTGGGCACTCTGCAACATTTTACACTGAAGAATGGCCAGTGGTC 1380
 DB |||||
 QY 1381 ATGCTGTGCTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 DB |||||
 QY 1381 ATGCTGTGCTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 DB |||||
 QY 1441 GATTTGATAAAGAACACATGAGTATATTAACACCCAGGGTGGCTTGTGCAAGGGGGT 1500
 DB |||||
 QY 1441 GATTTGATAAAGAACACATGAGTATATTAACACCCAGGGTGGCTTGTGCAAGGGGGT 1500
 DB |||||
 QY 1501 TAGGGCCATAGCAGTGTGTTACGACCAATAGACACAGGGCCCTATACGTTTATGTTGCTAC 1560
 DB |||||
 QY 1501 TAGGGCCATAGCAGTGTGTTACGACCAATAGACACAGGGCCCTATACGTTTATGTTGCTAC 1560
 DB |||||
 QY 1561 AAGGCTTTTCAAGTCCCAATTAAGTACCGGCTTGACAGATGATCTTACCGATATGATGGAT 1620
 DB |||||
 QY 1561 AAGGCTTTTCAAGTCCCAATTAAGTACCGGCTTGACAGATGATCTTACCGATATGATGGAT 1620
 DB |||||
 QY 1621 ACCCAGATGTGACCACTCTTAAAGGACAGCCGATTTTCCGTTTACTTGTGACACAGCTGTG 1680
 DB |||||
 QY 1621 ACCCAGATGTGACCACTCTTAAAGGACAGCCGATTTTCCGTTTACTTGTGACACAGCTGTG 1680
 DB |||||
 QY 1681 ATAGTCAGTGGAAACCACTGCTGTTTGGGGGAAACACACAAATGACACATCTATGAGC 1740
 DB |||||
 QY 1681 ATAGTCAGTGGAAACCACTGCTGTTTGGGGGAAACACACAAATGACACATCTATGAGC 1740
 DB |||||
 QY 1741 CATGGCCCAATGCTTCTTTCAGATTTTATGAGGCTATGACATTTGCTGTCACCGCTGG 1800
 DB |||||
 QY 1741 CATGGCCCAATGCTTCTTTCAGATTTTATGAGGCTATGACATTTGCTGTCACCGCTGG 1800
 DB |||||
 QY 1801 TCAGTCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCACGATC 1860
 DB |||||
 QY 1801 TCAGTCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCACGATC 1860
 DB |||||
 QY 1861 TTACACAAACAGCACCATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1920
 DB |||||
 QY 1861 TTACACAAACAGCACCATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1920
 DB |||||
 QY 1921 CTGTTATTTCACTCGGAACAGTGTGATGGGCAATCGAGTGAAGCGCTTGTGTAGCAGA 1980

FT /tag= a
 PN /product= "Membrane attractin-2"
 PD WO200015651-A1.

XX PN
 XX PD
 XX PF
 XX PR
 XX PA
 XX PI

DR 14-SEP-1999; 99WO-US020948.

XX 14-SEP-1998; 98US-0100137P.

XX (DAND) DANA FARBER CANCER INST INC.

XX Duke-Cohan JS, Schlossman SF;

XX WPI; 2000-271373/23.

XX P-PSDB; AAY70691.

PT Isolated nucleic acids encoding human attractin polypeptides useful for
 PT enhancing immune responses.

XX Claim 3; Fig 14; 120pp; English.

XX The patent discloses four forms of human attractin polypeptides which
 CC enhance immune response by promoting macrophage and monocyte spreading in
 CC the presence of T cells. These include soluble attractin-1 and -2 and
 CC membrane attractin-1 and -2. These various forms of attractin are encoded
 CC by alternatively spliced mRNA molecule transcribed from a single gene.
 CC The present sequence is a cDNA encoding human membrane attractin-2.
 CC Membrane attractin differs from soluble attractin in having a
 CC transmembrane domain and a cytoplasmic domain at the C-terminal.
 CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
 CC the N-terminal. This sequence can be used to enhance immune response in
 CC immunosuppressed patients such as those undergoing chemo- and radio-
 CC therapy treatment for cancer or those suffering from common variable
 CC immunodeficiency syndrome. The proteins may also be used to screen
 CC modulators (agonists and antagonists) of immune response which may also
 CC be used to regulate immune reactions. Attractin antibodies can be used to
 CC inhibit immune response in transplant recipients or patients afflicted
 CC with autoimmune disease

XX SQ Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;

Query Match 99.5%; Score 3801.4; DB 3; Length 4290;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3802; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGGTGGCGCAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA	60
DB	1	ATGGTGGCGCAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA	60
QY	61	GGCTCGCGGCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	120
DB	61	GGCTCGCGGCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	120
QY	121	CGGGGCTGGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	180
DB	121	CGGGGCTGGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	180
QY	181	CTG	240
DB	181	CTG	240
QY	241	GCCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300
DB	241	GCCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300
QY	301	TGTGACCGGCGGCTGTGTCAACGGGCGGCTGTGTCAACCGTGTGTGTGTGTGTGTGTGTGTG	360
DB	301	TGTGACCGGCGGCTGTGTCAACGGGCGGCTGTGTCAACCGTGTGTGTGTGTGTGTGTGTGTG	360
QY	361	CCCGCGGCTGGGTGGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAAGCTGGA	420

DB	361	CCCGCGGCTGGGTGGGCGGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAAGCTGGA	420
QY	421	TCCTCTGGGTTTGTGACAGATGACCTGGGAAATATATATATATATATATATATATATATATAT	480
DB	421	TCCTCTGGGTTTGTGACAGATGACCTGGGAAATATATATATATATATATATATATATATATAT	480
QY	481	CTCATTCAGAGCAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTGCTACAGAG	540
DB	481	CTCATTCAGAGCAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTGCTACAGAG	540
QY	541	TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCT	600
DB	541	TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCT	600
QY	601	GCAATTTAGTGGCTCATTTGTTCTCTGAGAGATGGCAATGAGACTCTCCCTGAGGTTGTT	660
DB	601	GCAATTTAGTGGCTCATTTGTTCTCTGAGAGATGGCAATGAGACTCTCCCTGAGGTTGTT	660
QY	661	GCCACATCAGGTTATGCTTGTCTGCTGCAATTTTATGATGATGCTGCTTATAATTTGACTGA	720
DB	661	GCCACATCAGGTTATGCTTGTCTGCTGCAATTTTATGATGATGCTGCTTATAATTTGACTGA	720
QY	721	TTTAAATATTACTTACAGTTTGTGATATGTTCCAAATAACTGCTCAGGCGGAGAGTCT	780
DB	721	TTTAAATATTACTTACAGTTTGTGATATGTTCCAAATAACTGCTCAGGCGGAGAGTCT	780
QY	781	AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTGAA	840
DB	781	AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTGAA	840
QY	841	GCAATGACATTCCTCCTCAGTGTACAGCAACTGTTGTTTCTTCATCAGGCGATCTGCAT	900
DB	841	GCAATGACATTCCTCCTCAGTGTACAGCAACTGTTGTTTCTTCATCAGGCGATCTGCAT	900
QY	901	TCAAGTCAATGTCAGGAGTCTCTCTCTCTCAGACTGCGAGGTCCTGATGTTGAGTT	960
DB	901	TCAAGTCAATGTCAGGAGTCTCTCTCTCTCAGACTGCGAGGTCCTGATGTTGAGTT	960
QY	961	CTGTACCAGCTAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
DB	961	CTGTACCAGCTAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
QY	1021	AGAGCATCTCAATAAGCTGTGTCAATGCAACATTTATGCTGGTGTGTTGAGGATATATG	1080
DB	1021	AGAGCATCTCAATAAGCTGTGTCAATGCAACATTTATGCTGGTGTGTTGAGGATATATG	1080
QY	1081	TTCAACCACTCAGATTTAATCAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
DB	1081	TTCAACCACTCAGATTTAATCAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
QY	1141	CCACTTAAACCGTCTGTGCAACATGTTGTTAGATATGTTGCTGCTGCTGCTGCTGCTGCTG	1200
DB	1141	CCACTTAAACCGTCTGTGCAACATGTTGTTAGATATGTTGCTGCTGCTGCTGCTGCTGCTG	1200
QY	1201	AAGGATAAAATTTACATGTTATGAGGAGGAAAAATGATCCAACTGGGAAATGTGACCAATGAG	1260
DB	1201	AAGGATAAAATTTACATGTTATGAGGAGGAAAAATGATCCAACTGGGAAATGTGACCAATGAG	1260
QY	1261	TTGAGAGTTTTCATTTTCAATTAATGATCATGGTGTGTTGACCCCTTAAGGCAAGGAG	1320
DB	1261	TTGAGAGTTTTCATTTTCAATTAATGATCATGGTGTGTTGACCCCTTAAGGCAAGGAG	1320
QY	1321	CAGTATGAGTGGTGGGCACTCTGCACATTTTACACTGAAGAAATGGCGGAGTGTCT	1380
DB	1321	CAGTATGAGTGGTGGGCACTCTGCACATTTTACACTGAAGAAATGGCGGAGTGTCT	1380
QY	1381	ATGCTGGTCACTTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGAGGAATAT	1440
DB	1381	ATGCTGGTCACTTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGAGGAATAT	1440
QY	1441	GATTTGGATTAAGAACACATGAGGATATATATACACCCAGGGTGCCTTGTGCAAGGGGT	1500
DB	1441	GATTTGGATTAAGAACACATGAGGATATATATACACCCAGGGTGCCTTGTGCAAGGGGT	1500

Db 3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAGGAGTACAAAGATAGTTCTCTAAT 3720
 QY 3721 GAGAAGTTTGATTTTCGCAACACCAACAAATATACACTTTCTTTGTTTATGTACAGTAATTC 3780
 Db 3721 GAGAAGTTTGATTTTCGCAACACCAACAAATATACACTTTCTTTGTTTATGTACAGTAATTC 3780
 QY 3781 ACCTGCCCATCAAAATTCAGGT 3803
 Db 3781 ACCTGCCCATCAAAATTCAGAT 3803

RESULT 3

AAS72659
 ID AAS72659 standard; cDNA; 8561 BP.

XX AAS72659;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #8463.

XX KW Human; Chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG08472.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 1; SEQ ID NO 8463; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Query Match 99.5%; Score 3801.4; DB 5; Length 8561;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3802; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGTCGCGAGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGACGCGCGGCGAGCGGA 60
 Db 1 ATGTGTCGCGAGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGACGCGCGGCGAGCGGA 60
 QY 61 GCGCTCGCGGCGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Db 61 GCGCTCGCGGCGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 121 CCGGGGCTGGGGGCGGGGCTGCGGCTCCCGGCGGCTGCTGCTCCACCGCTGCGGCGGCGG 180
 Db 121 CCGGGGCTGGGGGCGGGGCTGCGGCTCCCGGCGGCTGCTGCTCCACCGCTGCGGCGGCGG 180
 QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 Db 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 QY 241 GCGGAGCG 300
 Db 241 GCGGAGCG 300
 QY 301 TGTGACCGGCGGCTGCTCAACGCGCGGTCGTCGCAACCTGCGGCGGCGGCGGCGGCGG 360
 Db 301 TGTGACCGGCGGCTGCTCAACGCGCGGTCGTCGCAACCTGCGGCGGCGGCGGCGGCGG 360
 QY 361 CCGCGCGGCTGGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 Db 361 CCGCGCGGCTGGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 QY 421 TCCTTCTGGGTTTGTGACAGATGACCTGGAAATTAATAATAACAAAGAGTGCAGTGG 480
 Db 421 TCCTTCTGGGTTTGTGACAGATGACCTGGAAATTAATAATAACAAAGAGTGCAGTGG 480
 QY 481 CTCATTGAGGACAGCCCAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAA 540
 Db 481 CTCATTGAGGACAGCCCAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAA 540
 QY 541 TGTAGTTGGGACCAATTAATATGTTATGATGGGAGTCAATTTATGACCGCTAGTGGT 600
 Db 541 TGTAGTTGGGACCAATTAATATGTTATGATGGGAGTCAATTTATGACCGCTAGTGGT 600
 QY 601 GCATTTAGTGGGCTCATTTGTTCTGAGAGAGATGGCAATGACACTGCTCCCTGAGGTGTT 660
 Db 601 GCATTTAGTGGGCTCATTTGTTCTGAGAGAGATGGCAATGACACTGCTCCCTGAGGTGTT 660
 QY 661 GCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 GCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 TTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 Db 721 TTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 QY 781 AAGATCAGTAATAGCAGGCAAACTGTTGAATGTTGAATGTTGAATGTTGAATGTTGAATGTT 840
 Db 781 AAGATCAGTAATAGCAGGCAAACTGTTGAATGTTGAATGTTGAATGTTGAATGTTGAATGTT 840
 QY 841 GCAATGACATTCCTCACTGTACAGCAAACTGTGGTTTTTCTCATCGAGGCATCTGCAAT 900
 Db 841 GCAATGACATTCCTCACTGTACAGCAAACTGTGGTTTTTCTCATCGAGGCATCTGCAAT 900
 QY 901 TCAAGTGATGTACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 901 TCAAGTGATGTACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 961 CTTGTACCAAGCTAACCAAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020
 Db 961 CTTGTACCAAGCTAACCAAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020

Db	2760	TTGCAAAACACAGTGTCTAAGCAGTGC	CGGACCACTGTGCTTTGAGGACAGCATGTGAGA	2819
Qy	2823	TTTGACACCGCGCAGCTCTGAGTGCAT	GTGGTGGCAGCAACATGAAGCAGTGTGTGGACTC	2882
Db	2820	TTTGACACCGCGCAGCTCTGAGTGCAT	GTGGTGGCAGCAACATGAAGCAGTGTGTGGACTC	2879
Qy	2883	CAATGCCCTATGTGGCCCTCTTCCCTTT	TGGCCAGTGTATGGAATGTATACGATGAGCAC	2942
Db	2880	CAATGCCCTATGTGGCCCTCTTCCCTTT	TGGCCAGTGTATGGAATGTATACGATGAGCAC	2939
Qy	2943	CTGCCCCCTCAAAATTTGTTCAAGGCT	ACTGTACCTGTAGTCTATTGCTTTGAGAGCAACCCAGG	3002
Db	2940	CTGCCCCCTCAAAATTTGTTCAAGGCT	ACTGTACCTGTAGTCTATTGCTTTGAGAGCAACCCAGG	2999
Qy	3003	CTGTGGCTGTGTACTGATGCCAGCAAT	ATCTGGCAAAAGGGAATGCAATAGAGGTTCCCTA	3062
Db	3000	CTGTGGCTGTGTACTGATGCCAGCAAT	ATCTGGCAAAAGGGAATGCAATAGAGGTTCCCTA	3059
Qy	3063	TAAAGGACCACTGGAAGTGCCTTCGCA	AGCCCTTACAGGAAATTTCTATCCACAGCCCT	3122
Db	3060	TAAAGGACCACTGGAAGTGCCTTCGCA	AGCCCTTACAGGAAATTTCTATCCACAGCCCT	3119
Qy	3123	GCTCAATTCAGCATGTGTCTAGAGGAC	GAGAGATACAATCTGTCTTTCAATCACTGTCC	3182
Db	3120	GCTCAATTCAGCATGTGTCTAGAGGAC	GAGAGATACAATCTGTCTTTCAATCACTGTCC	3179
Qy	3183	AGCTTGCCAAATGCAACGGCCACACGT	AAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGA	3242
Db	3180	AGCTTGCCAAATGCAACGGCCACACGT	AAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGA	3239
Qy	3243	GAACTGACCAAGCGCAAGCACTCGGAG	ACCTGCAATCTCTGGCTTCTACGGTGATCCCAC	3302
Db	3240	GAACTGACCAAGCGCAAGCACTCGGAG	ACCTGCAATCTCTGGCTTCTACGGTGATCCCAC	3299
Qy	3303	CAATGGAGGGAATGTCTGAGCCATGCA	AGTGCATGGCGCATCTCTGTGGCAACACCA	3362
Db	3300	CAATGGAGGGAATGTCTGAGCCATGCA	AGTGCATGGCGCATCTCTGTGGCAACACCA	3359
Qy	3363	CACGGGCAAGTGTCTCTGCACCAACCA	AGGCGCTCAAGGGGGAACGAGTGCAGCTATGTGA	3422
Db	3360	CACGGGCAAGTGTCTCTGCACCAACCA	AGGCGCTCAAGGGGGAACGAGTGCAGCTATGTGA	3419
Qy	3423	GGTAGAAAATCGATACCAAGGAACCTCT	CAGAGNAACATGTTATTATCTCTTCTTAT	3482
Db	3420	GGTAGAAAATCGATACCAAGGAACCTCT	CAGAGNAACATGTTATTATCTCTTCTTAT	3479
Qy	3483	TGACTATCAGTTTCACTTTAGTCTATCC	CAGGAAGATGATCGCTATTACACAGCTATCAA	3542
Db	3480	TGACTATCAGTTTCACTTTAGTCTATCC	CAGGAAGATGATCGCTATTACACAGCTATCAA	3539
Qy	3543	TTTTTGGCTACTCTCGACGAAACAAAC	ACAGGGAATTTGGACATGTTCAATGCCCTCCAA	3602
Db	3540	TTTTTGGCTACTCTCGACGAAACAAAC	ACAGGGAATTTGGACATGTTCAATGCCCTCCAA	3599
Qy	3603	GAAATTTCAACCTCAACATCACTCGGG	CTGCCAGTTTCTCAGCTGGAAACCCAGCGCTGGAGA	3662
Db	3600	GAAATTTCAACCTCAACATCACTCGGG	CTGCCAGTTTCTCAGCTGGAAACCCAGCGCTGGAGA	3659
Qy	3663	AGAGATGCCCTGTGTTTCAAAAACCAAC	ATTAAGGAGTACAAGATAGTTCCTCTAAATGA	3722
Db	3660	AGAGATGCCCTGTGTTTCAAAAACCAAC	ATTAAGGAGTACAAGATAGTTCCTCTAAATGA	3719
Qy	3723	GAAATTTGATTTTCGCAACCAACCAAT	ATCACTTTCTTTGTTTATGTCAAGTAATTTCCAC	3782
Db	3720	GAAATTTGATTTTCGCAACCAACCAAT	ATCACTTTCTTTGTTTATGTCAAGTAATTTCCAC	3779
Qy	3783	CTGGCCCATCAAAATTCAGGTGCAAACT	GAAACATGA 3819	
Db	3780	CTGGCCCATCAAAATTCAGGTGCAAACT	GAAACATGA 3816	

AAZ91920
ID AAZ91920 standard; cDNA; 8590 BP.
XX
XX
AC AAZ91920;
XX
XX
DT 08-JUN-2000 (first entry)
XX
XX
DE Human mahogany protein coding sequence #2.
XX
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
XX weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antibesity; antianorexic; anticachexic; ss.
XX
XX
OS Homo sapiens.
XX
XX
FN W0200005373-A2.
XX
XX
PD 03-FEB-2000.
XX
XX
XX 21-JUL-1999; 99WO-US016484.
XX
XX
PR 21-JUL-1998; 98US-0093630P.
XX
PR 20-OCT-1998; 98US-0104978P.
XX
PR 05-FEB-1999; 99US-00245041.
XX
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
XX
PI Moore K, Nagle DL;
XX
XX
DR WPI; 2000-195103/17.
XX
XX
DR P-PSDB; AAY81807.
XX
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
XX treatment of body weight disorders.
XX
PS Claim 1; Fig 18a; 188pp; English.
XX
XX
CC This sequence represents a human mahogany gene of the invention. The
XX mahogany genes are used: (i) to produce recombinant mahogany (mg)
XX proteins (ii); (ii) as a source of antisense, ribozyme or triplex-forming
XX therapeutics; (iii) as a source of diagnostic probes and primers for
XX detecting expression of mg genes or mutations, regulatory defects, in
XX this gene, or for isolation of related sequences; and (iv) in (cell-
XX based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
XX identify other (extra)cellular products involved in weight regulation,
XX and to screen for agents that disrupt interaction between (ii) and other
XX macromolecules. The Ab are used to detect abnormal levels (or function)
XX of (ii) (for diagnosis, prognosis or monitoring of treatment); to
XX evaluate (ii)-expressing cells intended for cell therapy, and as
XX therapeutic mg inhibitors. Cells that express the mg gene (or contain the
XX mg polypeptide) are used to identify agents (A) that modulate mg
XX activity. (A) are potentially useful for the treatment of body weight
XX disorders, particularly obesity, cachexia or anorexia, or other
XX conditions associated with the mg gene such as hyperpigmentation,
XX hyperphagia and disorders that result in increased metabolic rate
XX
XX
SQ Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;
XX

CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human membrane attractin-1.
CC Membrane attractin differs from soluble human membrane attractin in having a
CC transmembrane domain and a cytoplasmic domain at the C-terminal. This
CC sequence can be used to enhance immune response in immunosuppressed
CC patients such as those undergoing chemo- and radio-therapy treatment for
CC cancer or those suffering from common variable immunodeficiency syndrome.
CC The proteins may also be used to screen modulators (agonists and
CC antagonists) of immune response which may also be used to regulate immune
CC reactions. Attractin antibodies can be used to inhibit immune response in
CC transplant recipients or patients afflicted with autoimmune disease
XX
SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;
Query Match 91.48; Score 3492.4; DB 3; Length 4068;
Best Local Similarity 99.38; Pred. No. 0;
Matches 3508; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 270 GTCGGGCTCAGCCGACGCGGCTCGGGGACGAGGCGGCGGCGGCTGTGTCAACGCGGCTCG 329
DB 48 GCGCGGACGCGGCTCGGGGACGAGGCGGCGGCGGCGGCTGTGTCAACGCGGCTCG 107
QY 330 CTGCAACCCCTGGCACCGGCGGCTGCGGCGGCGGCTGGTGGGCGGCAATGCCA 389
DB 108 CTGCAACCCCTGGCACCGGCGGCTGCGGCGGCGGCTGGTGGGCGGCAATGCCA 167
QY 390 GCACTGGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTGTGACAGATGACCTGG 449
DB 168 GCACTGGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTGTGACAGATGACCTGG 227
QY 450 AAATTATTAATACAAACGAGTGCAGTGGCTCATTTGAGGACGAGCAATGAATAAT 509
DB 228 AAATTATTAATACAAACGAGTGCAGTGGCTCATTTGAGGACGAGCAATGAATAAT 287
QY 510 GAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGA 569
DB 288 GAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGA 347
QY 570 TGGGACTCAATTTATGACCGGCTAGTTGCTGCAATTTAGTGGGCTCAATTTCTCTGAGAG 629
DB 348 TGGGACTCAATTTATGACCGGCTAGTTGCTGCAATTTAGTGGGCTCAATTTCTCTGAGAG 407
QY 630 AGATGGCAATGAGACTGTCCTGAGTTGTCACATCAGTTATGCTTGTGTCATTT 689
DB 408 AGATGGCAATGAGACTGTCCTGAGTTGTCACATCAGTTATGCTTGTGTCATTT 467
QY 690 TTTTATGATGCTGCTTATAATTTGACTCGAATTTAATTAATTAATTTACAGTTTGTATG 749
DB 468 TTTTATGATGCTGCTTATAATTTGACTCGAATTTAATTAATTAATTTACAGTTTGTATG 527
QY 750 TCCAAATAACTGCTCAGGCGGAGGAGTGAAGATCAGTAATAGCAGGAACTGTTGA 809
DB 528 TCCAAATAACTGCTCAGGCGGAGGAGTGAAGATCAGTAATAGCAGGATCTGTTGA 587
QY 810 ATGTGAATGCTCTGAAACTGAAAGGTGAAGCATGTCATCTCCTACTGTACAGCAA 869
DB 588 ATGTGAATGCTCTGAAACTGAAAGGTGAAGCATGTCATCTCCTACTGTACAGCAA 647
QY 870 CTGTGGTTTCTCATFCAGGCACTGCGAATTCAGATGATGATGATGATGATGATGATG 929
DB 648 CTGTGGTTTCTCATFCAGGCACTGCGAATTCAGATGATGATGATGATGATGATGATG 707
QY 930 CTCAGACTGGCAGGTCCTGGATGTTTCACTTCTGTACCGATTAACCACTCATTTTGGAC 989
DB 708 CTCAGACTGGCAGGTCCTGGATGTTTCACTTCTGTACCGATTAACCACTCATTTTGGAC 767
QY 990 TCGAGAGGAATATTCCTAAAGCTCCCGAGGATCTCAAGCTGTGGTCAATGG 1049
DB 768 TCGAGAGGAATATTCCTAAAGCTCCCGAGGATCTCAAGCTGTGGTCAATGG 827
QY 1050 AAACATTAATGTTGGTGTGGAGGATATATGTTCAACCACTCAGATTATTAACATGTTCT 1109
DB 828 AAACATTAATGTTGGTGTGGAGGATATATGTTCAACCACTCAGATTATTAACATGTTCT 887

QY 1110 AGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAATGTGCT 1169
DB 888 AGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAATGTGCT 947
QY 1170 TCTTAGATATGGTCACTCTTTGGCATTTACAAGATAAAATTTACATGTATGGAGGAAA 1229
DB 948 TGTTAGATATGGTCACTCTTTGGCATTTACAAGATAAAATTTACATGTATGGAGGAAA 1007
QY 1230 AATTGATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTC 1289
DB 1008 AATTGATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTC 1067
QY 1290 ATGGGTGTTTGTGACCCCTAAGGCAAGAGCAGTATGAGTGGTGGGCACTCTGCACA 1349
DB 1068 ATGGGTGTTTGTGACCCCTAAGGCAAGAGCAGTATGAGTGGTGGGCACTCTGCACA 1127
QY 1350 CAATTGTTACACTGAAGAATGGCCGAGTGGTCAATGCTGCTGCTCATCTTTGGTCACTGCCCTCT 1409
DB 1128 CAATTGTTACACTGAAGAATGGCCGAGTGGTCAATGCTGCTGCTCATCTTTGGTCACTGCCCTCT 1187
QY 1410 CTATGGATATATAAGCAATGTGACGAATATGATTTGGATAAGAACACATGGAGTATATT 1469
DB 1188 CTATGGATATATAAGCAATGTGACGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
QY 1470 ACACACCCAGGCTGCTTGTGCAAGGGGTTACCGGCATAGCAGTGTGTTTACGACCATAG 1529
DB 1248 ACACACCCAGGCTGCTTGTGCAAGGGGTTACCGGCATAGCAGTGTGTTTACGACCATAG 1307
QY 1530 GACCAGGGCCCTATAGCTTCAATGAGTGGCTTCAAGGCTTTTCAAGCAATAAGTACGGCT 1589
DB 1308 GACCAGGGCCCTATAGCTTCAATGAGTGGCTTCAAGGCTTTTCAAGCAATAAGTACGGCT 1367
QY 1590 TGCAGATGATCTCTACCGATATGATGGATACCCAGATGTGACCATCTTTAAAGACAG 1649
DB 1368 TGCAGATGATCTCTACCGATATGATGGATACCCAGATGTGACCATCTTTAAAGACAG 1427
QY 1650 CCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGAGCAACCATCTGCTGTTGG 1709
DB 1428 CCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGAGCAACCATCTGCTGTTGG 1487
QY 1710 GGGAAACACACACATCTATGAGCCATGSCGCGCAAAATGCTTCTTTCAGATT 1769
DB 1488 GGGAAACACACACATCTATGAGCCATGSCGCGCAAAATGCTTCTTTCAGATT 1547
QY 1770 CATGGCTATGACATTCGCTGACCGCTGGTCAAGTCTTCCAGACCTGATCTCCACCA 1829
DB 1548 CATGGCTATGACATTCGCTGACCGCTGGTCAAGTCTTCCAGACCTGATCTCCACCA 1607
QY 1830 TGATGCAACAGATTTGGCCATTCAGAGTCTTACACACAGCACCATGTATGTGTCGG 1889
DB 1608 TGATGCAACAGATTTGGCCATTCAGAGTCTTACACACAGCACCATGTATGTGTCGG 1667
QY 1890 TGGTTTCAATAGTCTCTCTGAGCAGCATCTGTTGTTATTCACCTCGGAAACAGTGTGATG 1949
DB 1668 TGGTTTCAATAGTCTCTCTCTGAGCAGCATCTGTTGTTATTCACCTCGGAAACAGTGTGATG 1727
QY 1950 GCATCGAGTGAAGCCGCTTGTGTTAGCAGCAGGACCTGGTATTCGGTGTGTGGAACAC 2009
DB 1728 GCATCGAGTGAAGCCGCTTGTGTTAGCAGCAGGACCTGGTATTCGGTGTGTGGAACAC 1787
QY 2010 AGGTGCTCTCAGTGTATCTCTGCGGCTGGCAACTGATGAACAGAAAGATTA 2069
DB 1788 AGGTGCTCTCAGTGTATCTCTGCGGCTGGCAACTGATGAACAGAAAGATTA 1847
QY 2070 ATCAGATGTTTTCAGAAAGAACTCTTGTACCATGACAGATGTGACGACACAGATTG 2129
DB 1848 ATCAGATGTTTTCAGAAAGAACTCTTGTACCATGACAGATGTGACGACACAGATTG 1907
QY 2130 TTACAGCTGTACAGCCAAACCAATGATGCTGCCATGTGCAATGACCATTTGTGTCGCCAG 2189
DB 1908 TTACAGCTGTACAGCCAAACCAATGATGCTGCCATGTGCAATGACCATTTGTGTCGCCAG 1967

QY 2190 GAACACAGCTGCTCAGAGGCGCAGATCTCCATTTTATGATGAGATGCTCCCAAGGA 2249
Db 1968 GAACCAACAGCTGCTCAGAGGCGCAGATCTCCATTTTATGATGAGATGCTCCCAAGGA 2027
QY 2250 TAAACCCATGTTACTACTGTAACAAGAACACAGCTGCGAGAGCTGTGCCCTGGACAGAA 2309
Db 2028 TAAACCCATGTTACTACTGTAACAAGAACACAGCTGCGAGAGCTGTGCCCTGGACAGAA 2087
QY 2310 CTGCCAGTGGGAGCCCCGGAATCAGAGTGCATTCGCTCCCGGAAATATCTGTGGCAT 2369
Db 2088 CTGCCAGTGGGAGCCCCGGAATCAGAGTGCATTCGCTCCCGGAAATATCTGTGGCAT 2147
QY 2370 TGGCTGGCAATTTGGTTTGGAAATCTCATGTTTGAATAATTAATCTACTGCGCAAGGAGATTAATGA 2429
Db 2148 TGGCTGGCAATTTGGTTTGGAAATCTCATGTTTGAATAATTAATCTACTGCGCAAGGAGATTAATGA 2207
QY 2430 CAATGCTAAATTTGTTCTGTAGGAACCAATATGCCCTTTGGCTTCTCTTAACAACCCAGAA 2489
Db 2208 CAATGCTAAATTTGTTCTGTAGGAACCAATATGCCCTTTGGCTTCTCTTAACAACCCAGAA 2267
QY 2490 GAAGGTAGAAATTTGCTTAAAGCAGCTGCGAATATGAGTCAATCTCAGAGCATGTCCAA 2549
Db 2268 GAAGGTAGAAATTTGCTTAAAGCAGCTGCGAATATGAGTCAATCTCAGAGCATGTCCAA 2327
QY 2550 GGTCAACCTTAACCCCATGGGTGGGCTTTCCGAAGATCAATGTGCTCTACTGGTGGGA 2609
Db 2328 GGTCAACCTTAACCCCATGGGTGGGCTTTCCGAAGATCAATGTGCTCTACTGGTGGGA 2387
QY 2610 AGATATGCCCATTTACAAATAGTTTACTACAGTGGATGCGCTGAGCCAGTATGC 2669
Db 2388 AGATATGCCCATTTACAAATAGTTTACTACAGTGGATGCGCTGAGCCAGTATGC 2447
QY 2670 TGGATTTCTGTGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCAT 2729
Db 2448 TGGATTTCTGTGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCAT 2507
QY 2730 CAACCCATCAATGTAGTGTCTGTAAGAGGCTTGCAAAACACAGTGTCTAAGCAGTCCG 2789
Db 2508 CAACCCATCAATGTAGTGTCTGTAAGAGGCTTGCAAAACACAGTGTCTAAGCAGTCCG 2567
QY 2790 GACACCATGTGCTTGAGACAGCATGTGGAGATTGACACAGCGCGAGCTCTGAGTGCAT 2849
Db 2568 GACACCATGTGCTTGAGACAGCATGTGGAGATTGACACAGCGCGAGCTCTGAGTGCAT 2627
QY 2850 GTGGTGCAGCAACATGAAGCAGTGTGGAGTCCAAATGCCCTATGTGCCCTCTTCCCTTT 2909
Db 2628 GTGGTGCAGCAACATGAAGCAGTGTGGAGTCCAAATGCCCTATGTGCCCTCTTCCCTTT 2687
QY 2910 TGGCCAGTGTATGGAATGATACGATGAGCAGCTGCCCTGAAATTTGTTCCAGCTA 2969
Db 2698 TGGCCAGTGTATGGAATGATACGATGAGCAGCTGCCCTGAAATTTGTTCCAGCTA 2747
QY 2970 CTGTACCTGTAGTCAATTTGAGCAACACAGGCTGTGGCTGTGTACTGATCCAGCAA 3029
Db 2748 CTGTACCTGTAGTCAATTTGAGCAACACAGGCTGTGGCTGTGTACTGATCCAGCAA 2807
QY 3030 TACTGGCAAGGAAATGATAGAGGTTCTCTAATAAGAGCAGTGAAGATGCCCTTCGCA 3089
Db 2808 TACTGGCAAGGAAATGATAGAGGTTCTCTAATAAGAGCAGTGAAGATGCCCTTCGCA 2867
QY 3090 AGCCCTTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCAGCATGTGTCTAGAGGA 3149
Db 2868 AGCCCTTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCAGCATGTGTCTAGAGGA 2927
QY 3150 CAGCAGATACACTGGTCTTCTATCTCACTGTCCAGTTTGGCAATGCAACGGCCACAGTAA 3209
Db 2928 CAGCAGATACACTGGTCTTCTATCTCACTGTCCAGTTTGGCAATGCAACGGCCACAGTAA 2987
QY 3210 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTCAGCAGGCAAGCAGTCCGA 3269
Db 2988 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTCAGCAGGCAAGCAGTCCGA 3047
QY 3270 GACCTGCATATCTGGCTTCTACGGTGTATCCCAACAAATGGAGGGAATGTCTAGCCATGCAA 3329

Db 3048 GACCTGCATATCTGGCTTCTACGGTGTATCCCAATGGAGGAAATGTCTAGCCATGCAA 3107
QY 3330 GTGCAATGGGCAAGCGCTCTGTGCAACACCAACAGGCAAGTGTCTTGCAACCAACAA 3389
Db 3108 GTGCAATGGGCAAGCGCTCTGTGCAACACCAACAGGCAAGTGTCTTGCAACCAACAA 3167
QY 3390 GGGCGTCAAGGGGGAAGAGTGCAGTGTGAGGTAGAAAATCGATACCAAGGAAACCC 3449
Db 3168 GGGCGTCAAGGGGGAAGAGTGCAGTGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227
QY 3450 TCTCAGAGGAACATGTTATTTATATCTCTTATTGACTATCAGTTCACTTTAGTCTATC 3509
Db 3228 TCTCAGAGGAACATGTTATTTATATCTCTTATTGACTATCAGTTCACTTTAGTCTATC 3287
QY 3510 CCAGGAAGATGATCGCTATTATACACAGCTATCAATTTTGGCTACTCTCTGAGAAACAAA 3569
Db 3288 CCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGGCTACTCTCTGAGAAACAAA 3347
QY 3570 CAGGGATTTGGACATCTTCATCAATGCTCCAGANTTTCAACCTCAACATCACCTGGGC 3629
Db 3348 CAGGGATTTGGACATCTTCATCAATGCTCCAGANTTTCAACCTCAACATCACCTGGGC 3407
QY 3630 TGGCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAA 3689
Db 3408 TGGCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAA 3467
QY 3690 CATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTGATTTTGGCAACCAACCCAAA 3749
Db 3468 CATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTGATTTTGGCAACCAACCCAAA 3527
QY 3750 TATCACTTTCTTTGTTTATGTAGTCAATAATTTCACTGGCCCATCAAAATTCAGST 3803
Db 3528 TATCACTTTCTTTGTTTATGTAGTCAATAATTTCACTGGCCCATCAAAATTCAGAT 3581

RESULT 7

AAZ52274

ID AAZ52274 standard; cDNA; 3597 Bp.

XX AAZ52274;

XX AC

XX DT -18-JUL-2000 (first entry)

XX DE Human soluble attractin-1 cDNA.

XX KW Human; soluble attractin-1; immune response; macrophage; monocyte;

XX KW T cell; immunostimulant; immunosuppressed patient; cancer;

XX KW immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 1. 3597

XX FT /tag= a

XX FT /product= "Soluble attractin-1"

XX PN WO200015651-A1.

XX PD 23-MAR-2000.

XX PF 14-SEP-1999; 99WO-US020948.

XX PR 14-SEP-1998; 98US-0100137P.

XX PA (DAND) DANA PARBER CANCER INST INC.

XX PI Duke-Cohan JS, Schloeman SF;

XX DR WPI; 2000-271373/23.

XX DR P-PSDB; AAY70689.

XX PT Isolated nucleic acids encoding human attractin polypeptides useful for

PT	enhancing immune responses.	990	TCGAGAGGAATATTTCTAACTTAAGCTCCCGAGAGCATCTCATTAAGCTGTGGTCAATGG	1049
XX	Claim 3; Fig 8; 120pp; English.	768	TCGAGAGGAATATTTCTAACTTAAGCTCCCGAGAGCATCTCATTAAGCTGTGGTCAATGG	827
PS		1050	AAACATATATGTGGTGTGGAGGATATATTTCAACCACTCAGATTATAAATGTTCT	1109
XX	The patent discloses four forms of human attractin polypeptides which	828	AAACATATATGTGGTGTGGAGGATATATTTCAACCACTCAGATTATAAATGTTCT	887
CC	enhance immune response by promoting macrophage and monocyte spreading in	1110	AGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGTCTGTGAAAGTGGT	1169
CC	the presence of T cells. These include soluble attractin-1 and -2 and	888	AGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGTCTGTGAAAGTGGT	947
CC	membrane attractin-1 and -2. These various forms of attractin are encoded	1170	TGTTAGATATGTTTCACTTCTTGGCATTATACAAAGGATAAAATTTACATGTATGGAGAAA	1229
CC	by alternatively spliced mRNA molecule transcribed from a single gene.	948	TGTTAGATATGTTTCACTTCTTGGCATTATACAAAGGATAAAATTTACATGTATGGAGAAA	1007
CC	The present sequence is a cDNA encoding soluble attractin-1 obtained from	1230	AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTC	1289
CC	activated T cell and human fetal liver libraries. This sequence can be	1008	AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTC	1067
CC	used to enhance immune response in immunosuppressed patients such as	1290	ATGGGTGTTGTTGACCCCTTAAGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCACA	1349
CC	those undergoing chemo- and radio-therapy treatment for cancer or those	1068	ATGGGTGTTGTTGACCCCTTAAGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCACA	1127
CC	suffering from common variable immunodeficiency syndrome. The proteins	1350	CAITGTTTACACTGCAAGAAATGGCCGAGTGGTTCATGCTGGTCACTTGGTCACTGCCCTCT	1409
CC	may also be used to screen modulators (agonists and antagonists) of	1128	CAITGTTTACACTGCAAGAAATGGCCGAGTGGTTCATGCTGGTCACTTGGTCACTGCCCTCT	1187
CC	immune responses which may be used to regulate immune reactions.	1410	CTATGGATATATAGCAATGTGAGGAATATGATTTGGATAGAAACACATGGAGTATAT	1469
CC	Attractin antibodies can be used to inhibit immune response in transplant	1188	CTATGGATATATAGCAATGTGAGGAATATGATTTGGATAGAAACACATGGAGTATAT	1247
CC	recipients or patients afflicted with autoimmune disease	1470	ACACACCCAGGGTGGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTGTTACGACCATAG	1529
XX	Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;	1248	ACACACCCAGGGTGGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTGTTACGACCATAG	1307
SQ	Query Match 91.3%; Score 3487; DB 3; Length 3597;	1530	GACCAGGGCCCTATACGTTTCATGTTGGCTACAGGCTTTTCACTGCCAATTAAGTACCGGCT	1589
	Best Local Similarity 99.2%; Pred. No. 0;	1308	GACCAGGGCCCTATACGTTTCATGTTGGCTACAGGCTTTTCACTGCCAATTAAGTACCGGCT	1367
	Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;	1590	TGCAGATGATCTCTACCGATATGATGGATATCCAGATGTGGACCATCTCTTAAGGACAG	1649
QY	270 GTCCGGCTCAGCCGAGCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGTGC	1368	TGCAGATGATCTCTACCGATATGATGGATATCCAGATGTGGACCATCTCTTAAGGACAG	1427
DB	48 GCGCGCAGCGCAGCGCTCGCGGCGAGGCGCGCGGCGCACTGTGTCAACGGCGTGC	1650	CCGATTTTTCGTTTACCTTGACACAGCTGTGATAGTAGTGGAAACCAATGCTGGTGTGG	1709
QY	330 CTGCAACCTGACCGCGCAGTGTGCTGCCCCCGCGCTGGGTGGCGGAGCAATGCCA	1428	CCGATTTTTCGTTTACCTTGACACAGCTGTGATAGTAGTGGAAACCAATGCTGGTGTGG	1487
DB	108 CTGCAACCTGACCGCGCAGTGTGCTGCCCCCGCGCTGGGTGGCGGAGCAATGCCA	1710	GGGAAACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTCTCTTCAGATTT	1769
QY	390 GCACCTGCGGGCGCGCTTCAGACTAACTGGATTTCTGGTGTGTGAAGATGACCTGG	1488	GGGAAACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTCTCTTCAGATTT	1547
DB	168 GCACCTGCGGGCGCGCTTCAGACTAACTGGATTTCTGGTGTGTGAAGATGACCTGG	1770	CATGCCCTATGACATTCGCTGTGACCGCTGTGACCGCTGGTGGTCCAGACCTGATCTCCACCA	1829
QY	450 AAATATTAATAAACAAGCAAGTGCAGTGGCTCATTTGAAGCAGACGCAATAGATAAT	1548	CATGCCCTATGACATTCGCTGTGACCGCTGTGACCGCTGGTGGTCCAGACCTGATCTCCACCA	1606
DB	228 AAATATTAATAAACAAGTGCAGTGGCTCATTTGAAGCAGACGCAATAGATAAT	1830	TGATGTCAACAGATTTGGCCATT - CAGCAGTCTTACACACAGCAGCAGCATGTATGTTCG	1888
QY	510 GAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGACCACTTTATATGATGA	1607	TGATGTCAACAGATTTGGCCATT - CAGCAGTCTTACACACAGCAGCAGCATGTATGTTCG	1666
DB	288 GAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGACCACTTTATATGATGA	1889	GTGGTTTCAATAGTCTCTCTCTCAGCGACATCTCTGGTATTTCACTCCGAAACAGTGTGATG	1948
QY	570 TGGGACTCAATTTATGCAACCGCTAGTTCCTGATTTAGTGCCCTCATTTCTCTGAGAG	1667	GTGGTTTCAATAGTCTCTCTCTCAGCGACATCTCTGGTATTTCACTCCGAAACAGTGTGATG	1726
DB	348 TGGGACTCAATTTATGCAACCGCTAGTTCCTGATTTAGTGCCCTCATTTCTCTGAGAG	1949	CGCATCGGAGTGAAGCGCTTGTGTTAGCAGCAGGACCTGGTATTTCCGTTGTGTGGAAACA	2008
QY	630 AGATGGCAATGAGACTGCTCCCTGAGGTGTGTCACATCAGGTATGCTCTGCAATTT	1727	CGCATCGGAGTGAAGCGCTTGTGTTAGCAGCAGGACCTGGTATTTCCGTTGTGTGGAAACA	1786
DB	408 AGATGGCAATGAGACTGCTCCCTGAGGTGTGTCACATCAGGTATGCTTCTGCAATTT	2009	CAGGTCCTCTCAGTGTATCTCTGGGGCTGGGCACTGATGATGACAGAAAGATTTAA	2068
QY	690 TTTTACTGATGCTGCTTATATTTGACTCGATTTAATATTAATTACAGTTTTGATATGTG	1787	CAGGTCCTCTCAGTGTATCTCTGGGGCTGGGCACTGATGATGACAGAAAGATTTAA	1846
DB	468 TTTTACTGATGCTGCTTATATTTGACTCGATTTAATATTAATTACAGTTTTGATATGTG	2069	AATCAGATGTTTTCCTCAAAAAGACTCTTGACCATGACAGATGTGACCCAGCACACAGATT	2128
QY	750 TCCAAATTAATGCTCAGCGCGAGGAGTGAAGTCAATAGCAGGAACTGTGCA			
DB	528 TCCAAATTAATGCTCAGCGCGAGGAGTGAAGTCAATAGCAGGAACTGTGCA			
QY	810 ATGTGAATGTTCTGAAAACCTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA			
DB	588 ATGTGAATGTTCTGAAAACCTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA			
QY	870 CTGTGTTTTCCTCATGAGGCATCTGCAATTCAGTGTATGTCAGAGGATGCTCTGCTT			
DB	648 CTGTGTTTTCCTCATGAGGCATCTGCAATTCAGTGTATGTCAGAGGATGCTCTGCTT			
QY	930 CTCAGACTGCGGGCTCTGAGTTTCAGTTCTGTACAGCTTAACAGTCAATTTGGAC			
DB	708 CTCAGACTGCGGGCTCTGAGTTTCAGTTCTGTACAGCTTAACAGTCAATTTGGAC			

Drmanac RT, Liu C, Tang YT;
WPI; 2001-639362/73.
P-PSDB; ABG08470.
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
Claim 1; SEQ ID NO 8461; 103pp; English.
The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
coding sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 3597 BP; 951 A; 822 G; 898 G; 926 T; 0 U; 0 Other;
Query Match 91.3%; Score 3487; DB 5; Length 3597;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;
QY 270 GTCCGGCTCAGCGGAGCGAGCCAGGAATGTGACCGCCCTGTGTCACGGCGGTG 329
DB 48 GCGCGGAGCGGAGCGCTCGCGGAGGAGCGCGCGCGCACTGTGTCAACGGCGGTG 107
QY 330 CTGCAACCTCGGACCGGCGAGTGGCTCTGCCCGCGCGCTGGGTGGGCGAGCAATGCCA 389
DB 108 CTGCAACCTCGGACCGGCGAGTGGCTCTGCCCGCGCGCTGGGTGGGCGAGCAATGCCA 167
QY 390 GCACCTGCGGGGCGGCTTCAGACTAATCTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449
DB 168 GCACCTGCGGGGCGGCTTCAGACTAATCTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
QY 450 AATATATAATACAAACGAAGTGACGCGGTCTATTGAAGCACAGCCAAATAGAAATAAT 509
DB 228 AATATATAATACAAACGAAGTGACGCGGTCTATTGAAGCACAGCCAAATAGAAATAAT 287
QY 510 GAGACTTCGTTTCAATCAATTTTCTACAGAGTGTAGTGGGACCAATTTATATTTATGA 569
DB 288 GAGACTTCGTTTCAATCAATTTTCTACAGAGTGTAGTGGGACCAATTTATATTTATGA 347
QY 570 TGGGGACTCAATTTATGACCGGTAGTCTGCTGCTAGTGGCTATTAGTGGCTATTGTCCTGAGAG 629
DB 348 TGGGGACTCAATTTATGACCGGTAGTCTGCTGCTAGTGGCTATTAGTGGCTATTGTCCTGAGAG 407
QY 630 AGATGGCAATGAGACTGTCCCTGAGGTGTGTCACATCAGGTATTGCTTCTGCTCAATTT 689
DB 408 AGATGGCAATGAGACTGTCCCTGAGGTGTGTCACATCAGGTATTGCTTCTGCTCAATTT 467
QY 690 TTTTATGATGCTGCTTAAATTTGACTGGATTTAATATTACTACAGTTTATGATGTG 749
DB 468 TTTTATGATGCTGCTTAAATTTGACTGGATTTAATATTACTACAGTTTATGATGTG 527
QY 750 TCCAAATAACTGCTCAGGCGGAGGAGTGAAGATCAGTAATAGCAGCGAACTGTTGA 809

DB 528 TCCAAATAACTGCTCAGGCGGAGGAGTGAAGATCAGTAATAGCAGCGATCTTGA 587
QY 810 ATGTGAATGTTCTGAATACTGGAAGAGTGAAGCATGTGACATTCCTCACTGTACAGCAA 869
DB 588 ATGTGAATGTTCTGAATACTGGAAGAGTGAAGCATGTGACATTCCTCACTGTACAGCAA 647
QY 870 CTGTGGTTTCTCTCATCGAGGCATCTGCAATTCAGATGATGTCAGAGGATGCTCCTGCTT 929
DB 648 CTGTGGTTTCTCTCATCGAGGCATCTGCAATTCAGATGATGTCAGAGGATGCTCCTGCTT 707
QY 930 CTGAGACTGCGAGGCTCTGGATGTTTCACTTCTCTGACCAGCTAACCAAGTCATTTGGAC 989
DB 708 CTGAGACTGCGAGGCTCTGGATGTTTCACTTCTCTGACCAGCTAACCAAGTCATTTGGAC 767
QY 990 TCGAGAGGAATATCTTAAGCTCCCGAGCATCTCATAAAGCTGTGCTCAATGG 1049
DB 768 TCGAGAGGAATATCTTAAGCTCCCGAGCATCTCATAAAGCTGTGCTCAATGG 827
QY 1050 AATCATTTATGCTGGTGTGTTGGAGATATATGTTCAACCACTCAGATTATAACATGTTCT 1109
DB 828 AATCATTTATGCTGGTGTGTTGGAGATATATGTTCAACCACTCAGATTATAACATGTTCT 887
QY 1110 AGGTATGACTGCTTCTTAGGAGTGGCTTCCACTAAACCGTTCGTGAAACATGTGT 1169
DB 888 AGGTATGACTGCTTCTTAGGAGTGGCTTCCACTAAACCGTTCGTGAAACATGTGT 947
QY 1170 TGTATGATATGGTCAATCTTGTGATTAACAAGATATAATTTACATGTATGAGGAAA 1229
DB 948 TGTATGATATGGTCAATCTTGTGATTAACAAGATATAATTTACATGTATGAGGAAA 1007
QY 1230 AATTGATCCAACCTGGGAATGACCAATGATGATGAGTGTGAGATTTTTCATTTCAATGATGTC 1289
DB 1008 AATTGATCCAACCTGGGAATGACCAATGATGATGAGTGTGAGATTTTTCATTTCAATGATGTC 1067
QY 1290 ATGGGTGTTCTTGACCCCTAAGCAAGAGGAGTGTGAGTGTGGGCACTCTGCACA 1349
DB 1068 ATGGGTGTTCTTGACCCCTAAGCAAGAGGAGTGTGAGTGTGGGCACTCTGCACA 1127
QY 1350 CATTTGTACACTGAAGAATGCGCGAGTGTGCTGCTGCTCATCTTTGGTCACTGCCCTCT 1409
DB 1128 CATTTGTACACTGAAGAATGCGCGAGTGTGCTGCTGCTCATCTTTGGTCACTGCCCTCT 1187
QY 1410 CTATGATATATAAGCAATGTGCAAGATATGATTTGGATGAAGAACACATGGAGTATATT 1469
DB 1188 CTATGATATATAAGCAATGTGCAAGATATGATTTGGATGAAGAACACATGGAGTATATT 1247
QY 1470 ACACACCCAGGTCCTTGTGCAAGGGGTGTACGGCCATAGCAGTGTTTACGACCATAG 1529
DB 1248 ACACACCCAGGTCCTTGTGCAAGGGGTGTACGGCCATAGCAGTGTTTACGACCATAG 1307
QY 1530 GACCAGGGCCCTATACGTTTCACTGCTGCTACAAAGGCTTTTCAAGTCAATGATACCGCT 1589
DB 1308 GACCAGGGCCCTATACGTTTCACTGCTGCTACAAAGGCTTTTCAAGTCAATGATACCGCT 1367
QY 1590 TGCAGATGATCTTACCGATATGATGGATACCCAGATGGGACCACTCTTAAAGGACAG 1649
DB 1368 TGCAGATGATCTTACCGATATGATGGATACCCAGATGGGACCACTCTTAAAGGACAG 1427
QY 1650 CCGATTTTTCGTTTACCTTGCACACAGCTGTGATAGTGTGAGTGAACCAATGCTGTTGG 1709
DB 1428 CCGATTTTTCGTTTACCTTGCACACAGCTGTGATAGTGTGAGTGAACCAATGCTGTTGG 1487
QY 1710 GGGAAACACACAAATGACATCTATGAGCCATGCGGCCCAATGCTTCTTTCAGATTT 1769
DB 1488 GGGAAACACACAAATGACATCTATGAGCCATGCGGCCCAATGCTTCTTTCAGATTT 1547
QY 1770 CATGCTTATGACATTCGCTGTGACCGCTGTGATGCTTCCAGAGCTGATCTCCACCA 1829
DB 1548 CATGCTTATGACATTCGCTGTGACCGCTGTGATGCTTCCAGAGCTGAT-TCCACCA 1606
QY 1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTTACAAACAGCAGCCATGTATGTGTG 1888

Db 1607 TGAATGCAACAGATTTGGCCATTCACAGAGCTCTTACACACAGCACCATTGATGTGTTG 1666
Qy 1889 GTGGTTTCAATAGTCTCTCTCAGGACATCTCGTATTTACCTCGGAACAGTGTGATG 1948
Db 1667 GTGGTTTCAATAGTCTCTCTCAGGACATCTCGTATTTACCTCGGAACAGTGTGATG 1726
Qy 1949 CGCATCGGAGTGAAGCCGCTTGTGTTAGCAGCAGGACCTCGTATTCGGTGTGTGGAACA 2008
Db 1727 CGCATCGGAGTGAAGCCGCTTGTGTTAGCAGCAGGACCTCGTATTCGGTGTGTGGAACA 1786
Qy 2009 CAGGTCCTCTCAGTGTATCTCGTGGCGCTGCAACTGATGAACAAGAGAAAAGTTAA 2068
Db 1787 CAGGTCCTCTCAGTGTATCTCGTGGCGCTGCAACTGATGAACAAGAGAAAAGTTAA 1846
Qy 2069 AATCAGATGTTTTTCCAAAGAACTCTTGAACATGACAGATGTGACCAAGCACAGATT 2128
Db 1847 AATCAGATGTTTTTCCAAAGAACTCTTGAACATGACAGATGTGACCAAGCACAGATT 1906
Qy 2129 GTTACAGCTGTACAGCCACCAATGACTGCACTGGTCAATGACCAATGTGTCCCA 2188
Db 1907 GTTACAGCTGTACAGCCACCAATGACTGCACTGGTCAATGACCAATGTGTCCCA 1966
Qy 2189 GGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTTAGGTATGAGAAATTCGCCCAAGG 2248
Db 1967 GGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTTAGGTATGAGAAATTCGCCCAAGG 2026
Qy 2249 ATAAACCCATGTACTGTATCAAGAGACAGAGCTGAGAGCTGTGCCCTGGACACAGA 2308
Db 2027 ATAAACCCATGTACTGTATCAAGAGACAGAGCTGAGAGCTGTGCCCTGGACACAGA 2086
Qy 2309 ACTGCCAGTGGAGCCCGGATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGCA 2368
Db 2087 ACTGCCAGTGGAGCCCGGATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGCA 2146
Qy 2369 TTGGCTGGCATTTGGTTGGAATCTCATGTTTGAATAATTAATCTGCAAGGAAATATG 2428
Db 2147 TTGGCTGGCATTTGGTTGGAATCTCATGTTTGAATAATTAATCTGCAAGGAAATATG 2206
Qy 2429 ACAATGCTAAATTTGTTCTGGAACCAATGCGCTTTGGCTCTCTTCAACCCAGA 2488
Db 2207 ACAATGCTAAATTTGTTCTGGAACCAATGCGCTTTGGCTCTCTTCAACCCAGA 2266
Qy 2489 AGAAGTAGAATTTGTCCTTAAGCAGCTGCGAATATGCAATCATCTCAGAGCATGTCCA 2548
Db 2267 AGAAGTAGAATTTGTCCTTAAGCAGCTGCGAATATGCAATCATCTCAGAGCATGTCCA 2326
Qy 2549 AGCTCACCCTAACCCCATGGCTCGGCTTCGGAAGATCATGTGCTTACTGGTCTGGG 2608
Db 2327 AGCTCACCCTAACCCCATGGCTTCGGAAGATCAATGTGCTTACTGGTCTGGG 2386
Qy 2609 AAGATATGTCCTCATTAATAGTTTACTACAGTGCCTGCTCAGGCCAGTGTATG 2668
Db 2387 AAGATATGTCCTCATTAATAGTTTACTACAGTGCCTGCTCAGGCCAGTGTATG 2446
Qy 2669 CTGGATTCGTGGAATTTTATCAGAACCCAGTACTCGGGAATGAAAGCTGCAACCTGCA 2728
Db 2447 CTGGATTCGTGGAATTTTATCAGAACCCAGTACTCGGGAATGAAAGCTGCAACCTGCA 2506
Qy 2729 TCAACCCACTCAATGCTAGTGTCTGTGAAGCCCTGCAACCCAGAGCTTAAGCAGTGCC 2788
Db 2507 TCAACCCACTCAATGCTAGTGTCTGTGAAGCCCTGCAACCCAGAGCTTAAGCAGTGCC 2566
Qy 2789 GGACACCATGTGCTTGAAGACAGCATGTGGAGATTGCAACCGGCGAGCTCTGAGTGA 2848
Db 2567 GGACACCATGTGCTTGAAGACAGCATGTGGAGATTGCAACCGGCGAGCTCTGAGTGA 2626
Qy 2849 TGTGGTGCAGCAACATGAGCAGTGTGGACTCCAAATGCTGTGCTTCCCTT 2908
Db 2627 TGTGGTGCAGCAACATGAGCAGTGTGGACTCCAAATGCTGTGCTTCCCTT 2686
Qy 2909 TTGGCCAGTGTATGGAATGGTATGAGATGAGCACCTGCCCTGCCCTGAAATTTGTCAGGCT 2968
Db 2687 TTGGCCAGTGTATGGAATGGTATGAGATGAGCACCTGCCCTGCCCTGAAATTTGTCAGGCT 2746

Qy 2969 ACTGTACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCA 3028
Db 2747 ACTGTACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCA 2806
Qy 3029 ATACTGCAAGAGGAATGATAGAGGTTCTTATAAGAGACAGTGAAGATCCCTTCG 3088
Db 2807 ATACTGCAAGAGGAATGATAGAGGTTCTTATAAGAGACAGTGAAGATCCCTTCG 2866
Qy 3089 AAGCCCTACAGAAATTTCTATCCACAGCCCTGTCTCAATTCACAGATGTGCTAGAGG 3148
Db 2867 AAGCCCTACAGAAATTTCTATCCACAGCCCTGTCTCAATTCACAGATGTGCTAGAGG 2926
Qy 3149 ACAGCAGATCAACTGTGCTTTTCAATTCATCTGTCCAGTTGCCAATGCAACGGCCACAGTA 3208
Db 2927 ACAGCAGATCAACTGTGCTTTTCAATTCATCTGTCCAGTTGCCAATGCAACGGCCACAGTA 2986
Qy 3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACACAGGCAACACTTCG 3268
Db 2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACACAGGCAACACTTCG 3046
Qy 3269 AGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGAAATGTCAGCCATGA 3328
Db 3047 AGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGAAATGTCAGCCATGA 3106
Qy 3329 AGTGCANTGGCAGCGTCTCTGTGCAACACCAACAGGGCAAGTCTTCTGACCAACCA 3388
Db 3107 AGTGCANTGGCAGCGTCTCTGTGCAACACCAACAGGGCAAGTCTTCTGACCAACCA 3166
Qy 3389 AGGCGGTCAAGGGGACAGTGCAGTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448
Db 3167 AGGCGGTCAAGGGGACAGTGCAGTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226
Qy 3449 CTCTCAGAGAAATGTATTTATCTCTTCTTATGACTATCAGTTCACCTTTAGTCTAT 3508
Db 3227 CTCTCAGAGAAATGTATTTATCTCTTCTTATGACTATCAGTTCACCTTTAGTCTAT 3286
Qy 3509 CCAGGAAGATGATCGTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAA 3568
Db 3287 CCAGGAAGATGATCGTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAA 3346
Qy 3569 ACAGGATTTGGACATGTTTATCAATTCCTTCCAGAAATTTCAACCTCAACATCACCTGG 3628
Db 3347 ACAGGATTTGGACATGTTTATCAATTCCTTCCAGAAATTTCAACCTCAACATCACCTGG 3406
Qy 3629 CTGCCATTTCTCAGCTGGNACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCA 3688
Db 3407 CTGCCATTTCTCAGCTGGNACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCA 3466
Qy 3689 ACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGTATTTTCGCAACCCCAA 3748
Db 3467 ACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGTATTTTCGCAACCCCAA 3526
Qy 3749 ATATCATTCTTCTTGTATGTCAGTAATTTCACTGGGCCATCAAAATTCAGTGCAAA 3808
Db 3527 ATATCATTCTTCTTGTATGTCAGTAATTTCACTGGGCCATCAAAATTCAGTGCAAA 3586
Qy 3809 CTGAACAATGA 3819
Db 3587 CTGAACAATGA 3597

RESULT 9

ADD70997

ID ADD70997 standard; DNA; 3597 BP.

XX ADD70997;

AC

XX

Dr 15-JAN-2004 (first entry)

XX

DE Human attractin gene SEQ ID NO:1.

XX

KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;

KW cyrostatic; gene therapy; human; gene; ds.
XX Homo sapiens.
OS WO2003061564-A2.
XX 31-JUL-2003.
XX 20-DEC-2002; 2002WO-US040718.
XX 21-DEC-2001; 2001US-0341815P.
XX 31-DEC-2001; 2001US-0343185P.
XX (GENE-) GENE LOGIC INC.
XX (LGBI-) LG BIOMEDICAL INST.
XX Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;
XX WPI; 2003-663343/62.
XX Diagnosing liver cancer cells, useful for treating liver cancer.
XX associated with chronic hepatitis or cirrhosis comprises detecting the
XX level of expression in a tissue sample of one or more genes associated
XX with cancerous liver tissues.
XX Claim 1; SEQ ID NO 1; 176pp; English.
XX The present invention describes a method for diagnosing liver cancer
XX cells comprising detecting the level of expression in a tissue sample of
XX one or more genes given in the specification (see ABD70997 to ADD71105),
XX where differential expression of the genes is indicative of liver cancer.
XX Also described: (1) detecting the progression of liver cancer in a
XX patient; (2) monitoring the treatment of a patient with liver cancer; (3)
XX treating a patient with liver cancer; (4) typing a liver disease in a
XX patient; (5) detecting the presence or progression of liver cancer in a
XX patient with chronic hepatitis or cirrhosis; (6) differentiating liver
XX cancer related to chronic hepatitis from liver cancer related to
XX cirrhosis; (7) screening for an agent capable of modulating the onset or
XX progression of liver cancer; (8) a composition comprising at least two
XX oligonucleotides comprising a sequence that specifically hybridises to
XX any of the genes; (9) a solid support comprising the at least two
XX oligonucleotides; (10) a computer system comprising a database containing
XX information identifying the level in liver tissue of a set of genes; (11)
XX a method for using the computer system to present information identifying
XX the expression level in tissue or cell of any of the genes; and (12) a
XX therapeutic agent for slowing or halting the progression of liver cancer.
XX The methods are useful for treating liver cancer associated with chronic
XX hepatitis or cirrhosis. The present sequence represents a specifically
XX claimed human gene sequence which is used in the exemplification of the
XX present invention.
XX SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;
Query Match 91.3%; Score 3487; DB 9; Length 3597;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;
QY 270 GTCGGCTCAGCGGAGCGAGCGAGCAAGGAATGTACCGGCCCTGTGTCAACGGCGGTGCG 329
DB 48 GCGGGCGACGGCAGCGCTCGCGGGAGAGCGCGGGCGCGCACTGTGTCAACGGCGGTGCG 107
QY 330 CTGCAACCTTGGACCGCGCAGTGTGTGTGCGCGCGCGCTGGGTGGCGGAGCAATGCCA 389
DB 108 CTGCAACCTTGGACCGCGCAGTGTGTGTGCGCGCGCGCTGGGTGGCGGAGCAATGCCA 167
QY 390 GCATCTCGGGGGCGCTTCAGACTACTCGACTACTCGACTACTCGACTACTCGACTACTCG 449
DB 168 GCATCTCGGGGGCGCTTCAGACTACTCGACTACTCGACTACTCGACTACTCGACTACTCG 227
QY 450 AATATTATAATACAAACGAAGTGACGTGGCTCAATTGAAGGACAGCCAAATAGATAAT 509
DB 228 AATATTATAATACAAACGAAGTGACGTGGCTCAATTGAAGGACAGCCAAATAGATAAT 287

510 GAGACTTCGTTTCAATCAATTTTCTACAGAGTGTAGTTGGGACCACTTTATATATGTTATGA 569
288 GAGACTTCGTTTCAATCAATTTTCTACAGAGTGTAGTTGGGACCACTTTATATATGTTATGA 347
570 TGGGACTCAATTTATGACACCGCTAGTTCTGCTCAATTTAGTGGCCCTCATTTGTTCTCTGAGAG 629
348 TGGGACTCAATTTATGACACCGCTAGTTCTGCTCAATTTAGTGGCCCTCATTTGTTCTCTGAGAG 407
630 AGATGCAATGAGACTGTCCCTGAGTGTGTTGGCAGCATCAGGTATGCTTCTGCTGCAATTT 589
408 AGATGCAATGAGACTGTCCCTGAGTGTGTTGGCAGCATCAGGTATGCTTCTGCTGCAATTT 467
690 TTTTAGTGTGCTGCTTATAATTTGACTCGATTATATTTACTTACAGTTTTGATATGTG 749
468 TTTTAGTGTGCTGCTTATAATTTGACTCGATTATATTTACTTACAGTTTTGATATGTG 527
750 TCCAAATAACTGCTCAGGCGGAGAGTGTAGATCAGTAAATAGCAGGAGAACTGTTGA 809
528 TCCAAATAACTGCTCAGGCGGAGAGTGTAGATCAGTAAATAGCAGGAGAACTGTTGA 587
810 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869
588 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
870 CTGTGTTTTCCTCATCGAGGCATCTGCAATTCAGTGTAGTGTAGAGGATGCTCTGCTT 929
648 CTGTGTTTTCCTCATCGAGGCATCTGCAATTCAGTGTAGTGTAGAGGATGCTCTGCTT 707
930 CTGAGACTGCGAGGTCCTGGATGTTTCAGTTCCTGTACCAGCTTAACAGCTCAATTTGGAC 989
708 CTGAGACTGCGAGGTCCTGGATGTTTCAGTTCCTGTACCAGCTTAACAGCTCAATTTGGAC 767
990 TCGAGAGGAATATTTCTAACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGCTCAATGG 1049
768 TCGAGAGGAATATTTCTAACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGCTCAATGG 827
1050 AAACATATGTTGGTGTGTTGGAGATATATGTTCAACCACTCAGATTAACATGTTCT 1109
828 AAACATATGTTGGTGTGTTGGAGATATATGTTCAACCACTCAGATTAACATGTTCT 887
1110 AGCTATGACTTGTCTTAGGAGTGGCTTCCACTAAACCGTCTGTGAAACAATGTGT 1169
888 AGCTATGACTTGTCTTAGGAGTGGCTTCCACTAAACCGTCTGTGAAACAATGTGT 947
1170 TGTAGATATGTTCAATTTTGGCATTTATACAAAGGATAAAATTTACATGTATGGAGAA 1229
948 TGTAGATATGTTCAATTTTGGCATTTATACAAAGGATAAAATTTACATGTATGGAGAA 1007
1230 AATTGATCCAACTGGGATGTGACCAATGAGTTGAGAGTTTTTTCACATTCATATGAGTC 1289
1008 AATTGATCCAACTGGGATGTGACCAATGAGTTGAGAGTTTTTTCACATTCATATGAGTC 1067
1290 ATGGTGTGTTGTGACCCCTTAAGGCAAGGAGCAGTATGCAAGTGTGGGCACTCTGCACA 1349
1068 ATGGTGTGTTGTGACCCCTTAAGGCAAGGAGCAGTATGCAAGTGTGGGCACTCTGCACA 1127
1350 CATTTGTACTGAGAAATGCGGAGTGTGCTGCTGCTCATCTTTGGTCACTGCGCTCT 1409
1128 CATTTGTACTGAGAAATGCGGAGTGTGCTGCTGCTCATCTTTGGTCACTGCGCTCT 1187
1410 CTATGGATATATAGCAATGTGCAAGGAATATGATTTTGGATAAGAAACATCGAGTATATT 1469
1188 CTATGGATATATAGCAATGTGCAAGGAATATGATTTTGGATAAGAAACATCGAGTATATT 1247
1470 ACACACCCAGGTTGCCCTTGTGAAAGGGGTTACGGCCATAGCAGTGTTCACACCAATAG 1529
1248 ACACACCCAGGTTGCCCTTGTGAAAGGGGTTACGGCCATAGCAGTGTTCACACCAATAG 1307
1530 GACACGGCCCTTACCTTCACTGTTGCTTACAGGCTTTTTCAGTGCCCAATAGTACCGGCT 1589
1308 GACACGGCCCTTACCTTCACTGTTGCTTACAGGCTTTTTCAGTGCCCAATAGTACCGGCT 1367
1590 TGCAGATGATCTCTACCGGATATGATGTGGATATCCAGATGTGGACCAATCTTCTTAAGGACAG 1649

1368 TGCGATGATCTCTACCGATATGATGTGATACCCAGATGTGACCAATCTTTAAGACAG 1427
1650 CCGATTTTCCGTTACTTGCACACAGCTGTGATAGTAGTGAACCAATGCTGGTGTGG 1709
1428 CCGATTTTCCGTTACTTGCACACAGCTGTGATAGTAGTGAACCAATGCTGGTGTGG 1487
1710 GGGAAACACACAAATGACACACTATAGAGCCATGCGCCAAATGCTTCTCTCAGATTT 1769
1488 GGGAAACACACAAATGACACACTATAGAGCCATGCGCCAAATGCTTCTCTCAGATTT 1547
1770 CATGGCCTATGACATTCCTGTCGACCGCTGTGACCTGTGCTTCCAGACCTGATCTCCACA 1829
1548 CATGGCCTATGACATTCCTGTCGACCGCTGTGACCTGTGCTTCCAGACCTGAT-TCACACA 1606
1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTACACAAACAGCACCATGTATGTGTCG 1888
1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAACAGCACCATGTATGTGTCG 1666
1889 GTGGTTTCAATAGTCTCTCCTCAGGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1948
1667 GTGGTTTCAATAGTCTCTCCTCAGGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1726
1949 CGCATCGAGTGAAGCCGCTGTGTTAGCAGCAGGACCTGCTATTCGGTGTGTGGAACA 2008
1727 CGCATCGAGTGAAGCCGCTGTGTTAGCAGCAGGACCTGCTATTCGGTGTGTGGAACA 1786
2009 CAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAGAAAGAAAGTTAA 2068
1787 CAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAGAAAGAAAGTTAA 1846
2069 AATCAGATGTTTTCCAAAGAACTCTTGACCATGACAGATGTGACACAGCAGATTT 2128
1847 AATCAGATGTTTTCCAAAGAACTCTTGACCATGACAGATGTGACACAGCAGATTT 1906
2129 GTTACAGCTGTACAGCAACACCAATGACTGCGCACTGGTGCATGACCAATGTGTCCCCA 2188
1907 GTTACAGCTGTACAGCAACACCAATGACTGCGCACTGGTGCATGACCAATGTGTCCCCA 1966
2189 GGAACACAGCTGCTCAGNAGGCGAGATCTCCATTTTATAGTATGAGATTTGCCCAAG 2248
1967 GGAACACAGCTGCTCAGNAGGCGAGATCTCCATTTTATAGTATGAGATTTGCCCAAG 2026
2249 ATAAACCCATGACTACTGTAAACAGAGACAGCAGCTGCAGAGCTGTCCCTGACACAGA 2308
2027 ATAAACCCATGACTACTGTAAACAGAGACAGCAGCTGCAGAGCTGTCCCTGACACAGA 2086
2309 ACTGCCAGTGGAGCCCGGAATCAGGAGTGCAATGGCCCTGCCGGAATAATCTGTGCA 2368
2087 ACTGCCAGTGGAGCCCGGAATCAGGAGTGCAATGGCCCTGCCGGAATAATCTGTGCA 2146
2369 TTGGCTGGCATTTGGTGTGAACCTCATGTTGAAATTAATCTGCAAGGAGAAATATG 2428
2147 TTGGCTGGCATTTGGTGTGAACCTCATGTTGAAATTAATCTGCAAGGAGAAATATG 2206
2429 ACAATGCTAAATTTGTTCTGTAGGAACCAATGGCCCTTTGGCTTCTCTTACAAACCCAGA 2488
2207 ACAATGCTAAATTTGTTCTGTAGGAACCAATGGCCCTTTGGCTTCTCTTACAAACCCAGA 2266
2489 AGAAGGTAGATTTGCTCCTTAAGCAGCTGCGAATAATGACATCATCTCAGAGCATGTCCA 2548
2267 AGAAGGTAGATTTGCTCCTTAAGCAGCTGCGAATAATGACATCATCTCAGAGCATGTCCA 2326
2549 AGCTCACCTTAACCCCATGGGTGGCCCTTCGGAAGATCAATGTGTCTTACTGGTGGG 2608
2327 AGCTCACCTTAACCCCATGGGTGGCCCTTCGGAAGATCAATGTGTCTTACTGGTGGG 2386
2609 AAGATATGCCCCATTAACAAATAGTTTACTACAGTGAATGCCCTGTGAGCCAGTGTATG 2668
2387 AAGATATGCCCCATTAACAAATAGTTTACTACAGTGAATGCCCTGTGAGCCAGTGTATG 2446
2669 CTGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728

2447 CTGGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCA 2506
2729 TCAACCCACTCAATGGTAGTGTCTGTGAAGGCTCTCAAAACACACAGTGTCTAAGCAGTGCC 2788
2507 TCAACCCACTCAATGGTAGTGTCTGTGAAGGCTCTCAAAACACACAGTGTCTAAGCAGTGCC 2566
2789 GGACACCATGTGCTTGGAGACAGCATGTGGAGATTTGCAACAGCGCAGCTCTGAGTGCA 2848
2567 GGACACCATGTGCTTGGAGACAGCATGTGGAGATTTGCAACAGCGCAGCTCTGAGTGCA 2626
2849 TGTGCTGCACACATGAACAGTGTGAGCTTCCCAATGCCTATGTGGCTCTCTCCCTT 2908
2627 TGTGCTGCACACATGAACAGTGTGAGCTTCCCAATGCCTATGTGGCTCTCTCTCCCTT 2686
2909 TTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCCTGAAAAATTTGTCAGGCT 2968
2687 TTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCCTGAAAAATTTGTCAGGCT 2746
2969 ACTGTACCTGTAGTCAATGCTTGGAGAACAGGCTGTGGCTGGTGTACTGATCCAGCA 3028
2747 ACTGTACCTGTAGTCAATGCTTGGAGAACAGGCTGTGGCTGGTGTACTGATCCAGCA 2806
3029 ATACTGGCAAGGAAATGCATAGAGGTTCTTATAAGGACAGTGAAGATGCTCTCGC 3088
2807 ATACTGGCAAGGAAATGCATAGAGGTTCTTATAAGGACAGTGAAGATGCTCTCGC 2866
3089 AAGCCCTTACAGAAATTTCTATCCACGCCCCCTGCTCAATTCACGATGTGTATAGAG 3148
2867 AAGCCCTTACAGAAATTTCTATCCACGCCCCCTGCTCAATTCACGATGTGTATAGAG 2926
3149 ACAGCAGATACAACTGGTCTTTCAATTCAGTCCAGTTCGCAATGCAACAGGCGCCAGTA 3208
2927 ACAGCAGATACAACTGGTCTTTCAATTCAGTCCAGTTCGCAATGCAACAGGCGCCAGTA 2986
3209 AATGCAATCAATCAGACATCTGTGAGAAGTGTGAGAACCTGACCAAGGAGAAATGCTGG 3268
2987 AATGCAATCAATCAGACATCTGTGAGAAGTGTGAGAACCTGACCAAGGAGAAATGCTGG 3046
3269 AGACCTGCATATCTGGCTTCTAGCGTGTATCCCAACCAATGGAGGAAATGCTGACCATGCA 3328
3047 AGACCTGCATATCTGGCTTCTAGCGTGTATCCCAACCAATGGAGGAAATGCTGACCATGCA 3106
3329 AGTGCAATGGCAGCGCTCTCTGTGCAACACCAACAGCGGCAAGTCTTCTGCAACCA 3388
3107 AGTGCAATGGCAGCGCTCTCTGTGCAACACCAACAGCGGCAAGTCTTCTGCAACCA 3166
3389 AGGCGCTCAAGGGGACGATGCGCAGCTATGTAGTAGAATAATCGATHACAGGAACCC 3448
3167 AGGCGCTCAAGGGGACGATGCGCAGCTATGTAGTAGAATAATCGATHACAGGAACCC 3226
3449 CTCTCAGAGAAACATGTTATATATCTCTTATTGACATCATGTTACCTTTAGTCTAT 3508
3227 CTCTCAGAGAAACATGTTATATATCTCTTATTGACATCATGTTACCTTTAGTCTAT 3286
3509 CCAGGAAGATGATGCTATTAACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAA 3568
3287 CCAGGAAGATGATGCTATTAACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAA 3346
3569 ACAGGATTTGGACATGTTTCATCAATGCTCCCAAGATTTCAACCTCAACATCAGCTGGG 3628
3347 ACAGGATTTGGACATGTTTCATCAATGCTCCCAAGATTTCAACCTCAACATCAGCTGGG 3406
3629 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCTCTGTTTCAAAAACCA 3688
3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCTCTGTTTCAAAAACCA 3466
3689 ACATTAAGAGTACAAAGATAGTTTCTTAATGAGAGTTTGAATTTTGGACACACCCAA 3748
3467 ACATTAAGAGTACAAAGATAGTTTCTTAATGAGAGTTTGAATTTTGGACACACCCAA 3526
3749 ATATCACTTTCTTTGTTTATGTAGTAAATTTCACTGGGCCCATCAAAATTCAGGTGCAAA 3808
3527 ATATCACTTTCTTTGTTTATGTAGTAAATTTCACTGGGCCCATCAAAATTCAGGTGCAAA 3586

QY 963 TGTACAGCTAACAGTCAATTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCCAG 1022
Db 1032 TGTGCCAGCTAACAGTCTTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCCAG 1091
QY 1023 AGCATCTCATAAGCTGTGGTCAATGGAAACATTTATGTGGTGTGGTGGAGATATATGTT 1082
Db 1092 AGCTCTCATAAGCTGTGGTCAATGGAAACATTTATGTGGTGTGGTGGAGATATATGTT 1151
QY 1083 CAACCACTCAGATTATAACATGGTCTAGCGATGACCTTGTCTTAGGGATGGCTTCC 1142
Db 1152 CAACCACTCAGATTATAACATGGTCTAGCGATGACCTTGTCTTAGGGATGGCTTCC 1211
QY 1143 ACTAAACCCCTCTGTGAAACAAATGTGGTGTAGATATGATGATCTTTGGCATATACAA 1202
Db 1212 ACTAAACCAATTCTGTGAACAGTGTGGTGTAGATATGATGATCTTTGGCATATACAA 1271
QY 1203 GGATAAAATTTACATGATGTGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1262
Db 1272 GGATAAAATTTACATGATGTGAGGAAATTTGATCCAACTGGGAATGTGACCAAGAGCT 1331
QY 1263 GAGAGTTTTTACATTCATAATGAGTCAATGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 1322
Db 1332 GAGAGTTTTTACATTCATAATGAGTCAATGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 1391
QY 1323 GTATGAGTGTGGTGGTCACTCTGACACATTTGTTACACTGAAGATGGCCGAGTGGTCA 1382
Db 1392 GTATGAGTGTGGTGGTCACTCTGACACATTTGTTACACTGAAGATGGCCGAGTGGTCA 1451
QY 1383 GCTGTGATCTTTGCTCACTGCTCTCTATGGATATATAGCAATGTGCAAGATATGA 1442
Db 1452 GTTGTGATCTTTGCTCACTGCTCTCTATGGATATATAGCAATGTGCAAGATATGA 1511
QY 1443 TTTGGATAAGAACACATGGAGTATATTAACACACAGGCTGCTTTGCAAGGGGGTTA 1502
Db 1512 CTTGGAAAGAACACATGGAGTATATTAACAACTCAAGTGTCTTGTGCAAGGGGGTTA 1571
QY 1503 CGGCCATAGCAGTGTGTAGCACTAGGACAGGCGCTTATAGTGTATGTGTGCTACAA 1562
Db 1572 TGGCCACAGCAGTGTGTATGATCAGGACAGGCGCTTATAGTGTGTGTGCTACAA 1631
QY 1563 GCTTTTCACTGCAATAGTACCGCTTGCAGATGATCTTACCGATATGATGTGGATAC 1622
Db 1632 GCTTTTCACTGCAATAGTACCGCTTGCAGATGATCTTACCGATATGATGTGGATAC 1691
QY 1623 CCAGATGTGGACCATTTCTTAAGACAGCGGATTTTCCGTTTACTTGCACACAGCTGTGAT 1682
Db 1692 TCAGATGTGGACCATTTCTTAAGACAGCGGATTTTCCGTTTACTTGCACACAGCTGTGAT 1751
QY 1683 AGTGAGTGGACCATGCTGTTTGGGGAAACACACACATGACATCTATGAGCCA 1742
Db 1752 AGTGAGTGGACCATGCTGTTTGGGGAAACACACACATGACATCTATGAGCCA 1811
QY 1743 TGGCCCAAAATGCTTCTCTTCAAGATTTTCAAGCTTATGACATTTGCTTGTGACCGCTGTC 1802
Db 1812 CGGTGCCAAATGCTTCTCTTCAAGATTTTCAAGCTTATGACATTTGCTTGTGACCGCTGTC 1871
QY 1803 AGTGCTTCCGACCTGATCTTCCACCATGATGTCAACAGATTTGGCCATTTGACAGCTT 1862
Db 1872 AGTCTTCCGACCTGATCTTCCACCATGATGTCAACAGATTTGGCCATTTGACAGCTT 1931
QY 1863 ACACAAACAGCAGTATGTTGCTGGTGGTTCATAGTCTCTCTCCTCAGCGACATCCT 1922
Db 1932 GCACACAGCAGCAGTATGTTGCTGGCGGCTTCAACAGCTTCTCTCAGTACGCTT 1991
QY 1923 GGTATTCAGCTCGGAACAGTGTGATGCGCATCGAGTGAAGCGGTTGTTTACAGCAGG 1982
Db 1992 AGTCTTCACTCGGACAGTGTGATGACACCGCAGCAAGCTGCTTGTGTGGCAGCAGG 2051
QY 1983 ACCTGGATTCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTGCTGGCGCTGGC 2042
Db 2052 ACCTGGATTCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTGCTGGCGCTGGC 2111
QY 2043 AACTGATGAACAGAGAAAGTTAAATCAGATGTTTTCAAAAAGAACTCTTGGACCA 2102

Db 2112 AACTGAGAAACAGAGAAAGTTAAATCAGAAATGTTTCTCTPAAAAGAACCTTGACCA 2171
QY 2103 TGACAGATGTGACAGCACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACGTGCCA 2162
Db 2172 TGACAGATGTGACAGCACACAGATTTGTTACAGCTGTACAGCCAAATACCAATGACGTGCCA 2231
QY 2163 CTGGTCAATGACCAATGTGTGCTCCAGAGCACACAGCTGCTCAGAGGCGCAGATCTCCAT 2222
Db 2232 CTGGTCAATGACCAATGTGTGCTCCAGAGCACACAGCTGCTCAGAGGCGCAGATCTCCAT 2291
QY 2223 TTTTAGTATGAGAAATTTGCCCAAGGATAACCCCATGTCTACTCTGTAACAAGAAAGCCAG 2282
Db 2292 TGCCAGATGACCAATTTGCCCAAGGATAACCCCATGTCTACTCTGCAATAAGAAACCCAG 2351
QY 2283 CTGACAGAGCTGTGCCCTGGACCCAGAACTGCGAGTGGAGCCCCCGAATCAGAGTGCAT 2342
Db 2352 CTGACAGAGCTGTGCCCTGGACCCAGAAATTTGCCAGTGGAAACCTCGAAATCAAGAGTGCAT 2411
QY 2343 TGCCCTGCCCGAAATATCTGTGGCAATTTGGCTGGCAATCTCATGTTTGA 2402
Db 2412 CGCCCTGCCCGAAACATCTGTGGCAATTTGGCTGGCAATCTCATGTTTGA 2471
QY 2403 AATTACTTACTGCAAGGAAATTTGACAAATGCTAAATTTGTTCTGTAGAAACCAATGC 2462
Db 2472 AATCACTTACTGCTAAGGAAATTTATGACAAATGCTAAATTTGCTGTAGAAACCAATGC 2531
QY 2463 CTTTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTTAAACAGCTGCGAAT 2522
Db 2532 CTTTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTTAAACAGCTGCGAAT 2591
QY 2523 AATGCAATCATCTCAGAGCAGCTGCAAGCTCACTTTAACCCCATGGCTGGCTTGGAA 2582
Db 2592 AATGCAATCATCTCAGAGCAGCTGCAAGCTCACTTTAACCCCATGGCTGGCTTGGAA 2651
QY 2583 GATCAATGTGCTACTGTGCTGGAGAAATATGTCCTTAAACAGCTGCGAAT 2642
Db 2652 GATCAATGTGCTACTGTGCTGGAGAAATATGTCCTTAAACAGCTGCGAAT 2711
QY 2643 GTGGATGCTGCTGAGCCCAAGCTGATCTGATTTCTGTGAAATTTTATCAGAAACCCAGTAC 2702
Db 2712 GTGGATGCTGCTGAGCCCAAGCTGATCTGATTTCTGTGAAATTTTATCAGAAACCCAGTAC 2771
QY 2703 TCGGGAGCTGAGGCTGCAACCTGCAATCAACCACTCAATGATGATGCTGTGAAAGGCC 2762
Db 2772 TCGGGAGCTGAGGCTGCAACCTGCAATCAACCACTCAATGATGATGCTGTGAAAGGCC 2831
QY 2763 TGCAAAACCAAGCTGCTAAGCAGTGGCGGACACATGCTGCTTGGAGACAGCATGTGGAGA 2822
Db 2832 TGCAAAACCAAGCTGCTAAGCAGTGGCGGACACATGCTGCTTGGAGACAGCATGTGGAGA 2891
QY 2823 TTGCACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAAGCAGTGTGTGAGCTC 2882
Db 2892 GTGCACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAAGCAGTGTGTGAGCTC 2951
QY 2883 CAATGCTATGTCCTGCTTCCCTTTGGCCAGTGTATGAAATGATGATACCATGAGCAG 2942
Db 2952 CAATGCTATGTCCTGCTTCCCTTTGGCCAGTGTATGAAATGATGATACCATGAGCAG 3011
QY 2943 CTGCCCCCTGAAAATTTGTTGAGGCTTACTGTACCTGTAGTCTTGTGGAGCAACCCAGG 3002
Db 3012 CTGCCCCCTGAAAATTTGTTGAGGCTTACTGTACCTGTAGTCTTGTGGAGCAACCCAGG 3071
QY 3003 CTGTGCTGCTGATGATCTGAGCAATCTGCGGAAAGGGAATGATAGAGGTTCTTA 3062
Db 3072 CTGTGCTGCTGATGATCTGAGCAATCTGCGGAAAGGGAATGATAGAGGTTCTTA 3131
QY 3063 TAAAGGACAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATPCCACAGCCCT 3122
Db 3132 TAAAGGACAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATPCCACAGCCCT 3191
QY 3123 GCTCAATTCACAGATGCTGATAGAGACAGAGATACAACTGCTTCTTCACTGTGCC 3182

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;
Query Match 73.6%; Score 2810.2; DB 5; Length 3490;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 3237; Conservative 0; Mismatches 53; Indels 145; Gaps 25;
QY 409 AGACTAACTGGATCTTCTGGGTTTGGACAGATGACCTGGAAATATATAATACAAAACG 468
DB 22 AGACTAACTGGATCTTCTGGGTTTGGACAGATGACCTGGAAATATATAATACAAAACG 81
QY 469 AAGTGACAGTGGCTCAATCAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCAT 528
DB 82 AAGTGACAGTGGCTCAATCAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCAT 141
QY 529 TTGTCTACAGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCA 588
DB 142 TTGTCTACAGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCA 201
QY 589 CGCTAGTTGCTGCATTTAGTGGCTCATGTTCTGAGAGATGGCAATGAGACTGTC 648
DB 202 CGCTAGTTGCTGCATTTAGTGGCTCATGTTCTGAGAGATGGCAATGAGACTGTC 261
QY 649 CTTGAGGTTGTTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTAT 708
DB 262 CTTGAGGTTGTTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTAT 321
QY 709 AATTGACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 768
DB 322 AATTGACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 381
QY 769 CGAGGAGAGTGAAGATCAGTAATAGCAGGAACTGTTGAATGTGAATGTCTCAAAAC 828
DB 382 CGAGGAGAGTGAAGATCAGTAATAGCAGGATCTGTTGAATGTGAATGTCTCAAAAC 441
QY 829 TGGAAAGTGAAGCATGTGCAATCTCTACTGTACAGCAACTGTGGTTTCTCATCGA 888
DB 442 TGGAAAGTGAAGCATGTGCAATCTCTACTGTACAGCAACTGTGGTTTCTCATCGA 501
QY 889 GGCATCTGCAATTCAGTGTGATGTCAGAGATGCTCTGCTCTCAGATGGCAGGCTCT 948
DB 502 GGCATCTGCAATTCAGTGTGATGTCAGAGATGCTCTGCTCTCAGATGGCAGGCTCT 561
QY 949 GGATGTTCACTGTTACAGCTAACAGTCAATTTT-GGACTCGAGAGGAATATCTTAA 1007
DB 562 GGATGTTCACTGTTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATATCTTAA 621
QY 1008 CTT--AAAGCTCCGAGACATCTCATAAGCTGTGTTCAATGGAAACATTAATGTGGGT 1065
DB 622 CTTTAAAGCTCCGAGACATCTCATAAGCTGTGTTCAATGGAAACATTAATGTGGGT 681
QY 1066 GTTGAGGATATATGTTCAACCACTCAGATTATAACATGTTTCTAGCGTATGACCTTCT 1125
DB 682 GTTGAGGATATATGTTCAACCACTCAGATTATAACATGTTTCTAGCGTATGACCTTCT 741
QY 1126 TCTAGGAGTGGCTTCCATTAACCGTCTGTGTGAACATGTTGTTGATATGTTGAT 1185
DB 742 TCTAGGAGTGGCTTCCATTAACCGTCTGTGTGAACATGTTGTTGATATGTTGAT 801
QY 1186 TCTTTGGCATTAACAGGATAAATTTTACATGTATGGAGGAAATTTGATCCAACTGGG 1245
DB 802 TCTTTGGCATTAACAGGATAAATTTTACATGTATGGAGGAAATTTGATCCAACTGGG 861
QY 1246 AATGTGAACAAATGAGTTGAGATTTTTCATTCATTAATAGTGTATGTTGTTGATG 1305
DB 862 AATGTGAACAAATGAGTTGAGATTTTTCATTCATTAATAGTGTATGTTGTTGATG 921
QY 1306 CTTAGGCAAGAGGACGATGATGTTGGTGGGCTCTGCAACATTTGTACACTGAAG 1365
DB 922 CTTAGGCAAGAGGACGATGATGTTGGTGGGCTCTGCAACATTTGTACACTGAAG 981
QY 1366 AATGCCGAGTGGTCAATGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATAAAGC 1425

DB 982 AATGCCGAGTGGTCAATGCTGGTCACTTTTGGTCACTGCCCTCTCTATGGATATAAGC 1041
QY 1426 AATGTGAGGAATATGATTTGGATAAGAACACATGAGTATATATACACCCAGGTGCC 1485
DB 1042 AATGTGAGGAATATGATTTGGATAAGAACACATGAGTATATATACACACCCAGGTGCC 1101
QY 1486 CTTGTGCAAGGGGTTACGGCCATACAGTGTGTTAGACCATAGGACCGGGCCCTATAC 1545
DB 1102 CTTGTGCAAGGGGTTACGGCCATACAGTGTGTTAGACCATAGGACCGGGCCCTATAC 1161
QY 1546 GTTCAATGTTGGCTACAGGCTTTTCAAGTCCAAATAGTACCGGCTTCAGATGATCTCTAC 1605
DB 1162 GTTCAATGTTGGCTACAGGCTTTTCAAGTCCAAATAGTACCGGCTTCAGATGATCTCTAC 1221
QY 1606 CGATATGATGTGGATACCCAGATGTGGACCATTTCTTAAAGACAGCCGATTTTCCGTTAC 1665
DB 1222 CGATATGATGTGGATACCCAGATGTGGACCATTTCTTAAAGACAGCCGATTTTCCGTTAC 1281
QY 1666 TTGCACACAGCTGTGATGAGTGGAAACCATGCTGGTGTGTTGGGGGAAACACACAAT 1725
DB 1282 TTGCACACAGCTGTGATGAGTGGAAACCATGCTGGTGTGTTGGAGAAACACACAAT 1341
QY 1726 GACATCTATAGCCATGCGCCAAATGCTTCTTCAAGATTTCAATGAGCTATGATCTC 1785
DB 1342 GACATCTATAGCCATGCGCCAAATGCTTCTTCAAGATTTCAATGAGCTATGATCTC 1401
QY 1786 GCCTGTGACCGCTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACATTT 1845
DB 1402 GCCTGTGACCGCTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACATTT 1461
QY 1846 GGCATTTACAGCTCTTACACACAGCACCATGATGTGTTGGTGTGTTTCAATGATCTC 1905
DB 1462 GGCATTTACAGCTCTTACACACAGCACCATGATGTGTTGGTGTGTTTCAATGATCTC 1521
QY 1906 CTCTCTGAGCAGCATCTGTTGATTTTCACTCGAACAGTGTGATGCGCATCGAGTGAAGCC 1965
DB 1522 CTCTCTGAGCAGCATCTGTTGATTTTCACTCGAACAGTGTGATGCGCATCGAGTGAAGCC 1581
QY 1966 GCTTGTGTTAGCAGCAGGACCTGTTTGGTGTGTTGGAAACACAGGCTGCTCAGTGT 2025
DB 1582 GCTTGTGTTAGCAGCAGGACCTGTTTGGTGTGTTGGTGTGTTTCAATGATCTC 1614
QY 2026 ATCTGTGGGCGTGGCACTGATGAACAGAAAGAAAGTAAATCAGAAATGTTTTC 2085
DB 1615 ----- 1614
QY 2086 AAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATGTTTACAGCTGTACAGCC 2145
DB 1515 -----CTTTGACCATGACAGATGTGACAGCACACAGATGTTTACAGCTGTACAGCC 1667
QY 2146 AACACCAATGATGTCACCTGTTGCAATGACATGTTGTTCCCGAGGAAACACAGCTGCTCA 2205
DB 1668 AACACCAATGATGTCACCTGTTGCAATGACATGTTGTTCCCGAGGAAACACAGCTGCTCA 1727
QY 2206 GAGGCGCAGATCTCCATTTTATGATGAAATGTTGCCCAAGGATAAACCCATGATCTAC 2265
DB 1728 GAGGCGC-----AGGTATGAAATGTTGCCCAAGGATAAACCCATGATCTAC 1773
QY 2266 TGTAAACAGAAAGACACAGCTGTCAGGAGCTGTGCGCTCGACAGAACTGCCAGTGGGAGCC 2325
DB 1774 TGTAAACAGAAAGACACAGCTGTCAGGAGCTGTGCGCTCGACAGAACTGCCAGTGGGAGCC 1833
QY 2326 CGGAATCAGAGTGCATGCTGCGGAAATATCTGTGGCATTTGGCTGGCATTTGGT 2385
DB 1834 CGGAATCAGAGTGCATGCTGCGGAAATATCTGTGGCATTTGGCTGGCATTTGGT 1893
QY 2386 GGAATCTCATGTTTGAATTTACTTCTGCGCAAGAGAAATTTATGCAATGCTAAATTTGTTTC 2445
DB 1894 GGAATCTCATGTTTGAATTTACTTCTGCGCAAGAGAAATTTATGCAATGCTAAATTTGTTTC 1953
QY 2446 TGTAGGAACCAATGCTGCTTTGGTCTCTTTTACAAACCCAGAGAGTA-GAATTTGT 2504

CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
 CC mg polypeptide) are used to identify agents (A) that modulate mg
 CC activity. (A) are potentially useful for the treatment of body weight
 CC disorders, particularly obesity, cachexia or anorexia, or other
 CC conditions associated with the mg gene such as hyperpigmentation,
 CC hyperphagia and disorders that result in increased metabolic rate
 XX
 SQ Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

Query Match 58.9%; Score 2248.2; DB 3; Length 2625;
 Best Local Similarity 96.8%; Pred. No. 0;
 Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

QY	3	GGTGGCCGAGCGGCGGCACTGAGCAAGGCTGAGGAGGAGCGGGCGGACGGCAGC	62
Db	9	GGTGGCCGAGCGGCGGCACTGAGCAAGGCTGAGGAGGAGCGGGCGGACGGCAGC	68
QY	63	GCTCGGGCAGAGCGGCGGCGGCACTGAGGAGGAGCGGAGCGGAGCGGAGGCC	122
Db	69	GCTCGGGCAGAGCGGCGGCGGCACTGAGGAGGAGCGGAGCGGAGCGGAGGCC	128
QY	123	GGGGCTGGGGCGGGCTCGGCTCCGGGGCTGCTGCTCCACCGCTGCGGGCCACGGCT	182
Db	129	GGGACCGCGCGCGGCTGCTGCTCCGGGGCTGCTGCTCGGGGGCTGCGGGCGGCGGCC	188
QY	183	GCTGCTGCTGCTGCTGCTCCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGGC	242
Db	189	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGGC	239
QY	243	CGAGCGCGGCGGCGGCGGCGGCTGCGGGCTGAGCGGCGGCGGCGGCGGCGGCGG	302
Db	240	CGAGCGCGGCTGCGGCTGCGGCGGCGGCGGCTGCGGGCTGAGCGGCGGCGGCGG	299
QY	303	TGACCGGCGGCTGCTGCAAGCGGCTGCTGCAAGCGGCGGCGGCGGCGGCGGCGG	362
Db	300	TGACCGGCGGCTGCTGCAAGCGGCTGCTGCAAGCGGCGGCGGCGGCGGCGGCGG	359
QY	363	CGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	422
Db	360	CGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	419
QY	423	TCTGGGCTTGTGACAGATGACCTGGAATTAATAAATAAATAAATAAATAAATAA	482
Db	420	TCTGGGCTTGTGACAGATGACCTGGAATTAATAAATAAATAAATAAATAAATAA	479
QY	483	CATTGAAGGACAGCCAAATAGATAATGAGACTTCGTTTCAATCATTTTGTACAGAGTG	542
Db	480	CATTGAAGGACAGCCAAATAGATAATGAGACTTCGTTTCAATCATTTTGTACAGAGTG	539
QY	543	TAGTTGGGACCAATTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTGTGC	602
Db	540	TAGTTGGGACCAATTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTGTGC	599
QY	603	ATTTAGTGGCTCATTTGTTCTGAGAGAGATGCAATGAGACTGCTCCCTGAGGTTGTGC	662
Db	600	ATTTAGTGGCTCATTTGTTCTGAGAGAGATGCAATGAGACTGCTCCCTGAGGTTGTGC	659
QY	663	CACATCAGGTTATGCTTGTGCTATTTTATGATGCTGCTTATATTTGACTGGATT	722
Db	660	CACATCAGGTTATGCTTGTGCTATTTTATGATGCTGCTTATATTTGACTGGATT	719
QY	723	TAATATTACTTACAGTTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	782
Db	720	TAATATTACTTACAGTTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	779
QY	783	GATCAGTAATAGACGGAACCTGTTGAATGTGATGTTCTGAAACCTGGAAGAGCTCAAGC	842
Db	780	GATCAGTAATAGACGCGATCTGTTGAATGTGATGTTCTGAAACCTGGAAGAGCTCAAGC	839
QY	843	ATGTGACATTCCTCAGTACAGCAACTGTGGTTTTCCTCATCGAGGCAATCGCAATTC	902
Db	840	ATGTGACATTCCTCAGTACAGCAACTGTGGTTTTCCTCATCGAGGCAATTCGCAATTC	899

QY	903	AAGTGATGTCAGAGGATGCTCTCTCAGACTGGCAGGGCTCTGGATGTTCAAGTTCC	962
Db	900	AAGTGATGTCAGAGGATGCTCTCTCAGACTGGCAGGGCTCTGGATGTTCAAGTTCC	959
QY	963	TGTACAGCTAACCAAGTCATTTTGGACTCGAGAGGATATTTAACTTAAAGCTCCCCAG	1022
Db	960	TGTACAGCTAACCAAGTCATTTTGGACTCGAGAGGATATTTAACTTAAAGCTCCCCAG	1019
QY	1023	AGCATCTCATAAAGCTGTGGTCAATGAAACATTAATGTGGTTTGTGGAGGATATATGTT	1082
Db	1020	AGCATCTCATAAAGCTGTGGTCAATGAAACATTAATGTGGTTTGTGGAGGATATATGTT	1079
QY	1083	CAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGTCTTCTAGGAGTGGCTTCC	1142
Db	1080	CAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGTCTTCTAGGAGTGGCTTCC	1139
QY	1143	ACTAAACCGTTCGTGTAACAAATGTTGTTAGATATGTCATTTCTTGGCATTTACAA	1202
Db	1140	ACTAAACCGTTCGTGTAACAAATGTTGTTAGATATGTCATTTCTTGGCATTTACAA	1199
QY	1203	GGATAAAATTTACATGATGAGGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT	1262
Db	1200	GGATAAAATTTACATGATGAGGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT	1259
QY	1263	GAGAGTTTTTCATTAATGAGTCATGGTGTGTTGACCCCTAAGGCAAGGAGCA	1322
Db	1260	GAGAGTTTTTCATTAATGAGTCATGGTGTGTTGACCCCTAAGGCAAGGAGCA	1319
QY	1323	GTATGAGTGGTTGGGCACTCTGCAACATTTGTTACCTGAAAGAAATGGCCGAGTGTGAT	1382
Db	1320	GTATGAGTGGTTGGGCACTCTGCAACATTTGTTACCTGAAAGAAATGGCCGAGTGTGAT	1379
QY	1383	GCTGGTTCATTTTGGTCACTGCGCTCTCTATGGAATATAAGCAATGTGACGAATATGA	1442
Db	1380	GCTGGTTCATTTTGGTCACTGCGCTCTCTATGGAATATAAGCAATGTGACGAATATGA	1439
QY	1443	TTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGGCTTGTGCAAGGGGTTA	1502
Db	1440	TTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGGCTTGTGCAAGGGGTTA	1499
QY	1503	CGGCCATAGACGTTTACGACCATAGGACACAGGCGCTTATACGTTCAATGGTGGCTACAA	1562
Db	1500	CGGCCATAGACGTTTACGACCATAGGACACAGGCGCTTATACGTTCAATGGTGGCTACAA	1559
QY	1563	GGCTTTCAGTGGCAATAGTACCGGCTGACATGATCTCTACCGATATGATGGATAC	1622
Db	1560	GGCTTTCAGTGGCAATAGTACCGGCTGACATGATCTCTACCGATATGATGGATAC	1619
QY	1623	CCAGATGGGACCATTTTAAAGGACAGCGGATTTTTCGGTTACTTGCACAGCTGTGAT	1682
Db	1620	CCAGATGGGACCATTTTAAAGGACAGCGGATTTTTCGGTTACTTGCACAGCTGTGAT	1679
QY	1683	AGTGAGTGGAAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCA	1742
Db	1680	AGTGAGTGGAAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCA	1739
QY	1743	TGGCGCCAAATGCTTCTCTTCAGATTTTCATGGCTTATGACATTTGCTGTGACCGTGTGC	1802
Db	1740	TGGCGCCAAATGCTTCTCTTCAGATTTTCATGGCTTATGACATTTGCTGTGACCGTGTGC	1799
QY	1803	AGTGCTTCCAGACCTGATCTCCACCATGATGTGCAACAGATTTGGGCCATTGACAGTCTT	1862
Db	1800	AGTGCTTCCAGACCTGATCTCCACCATGATGTGCAACAGATTTGGGCCATTGACAGTCTT	1859
QY	1863	ACACAAACAGCACCATGATGTTGTTGGTGGTTCATTAAGTCTCTCTCCTCAGCGACATCCT	1922
Db	1860	ACACAAACAGCACCATGATGTTGTTGGTGGTTCATTAAGTCTCTCTCCTCAGCGACATCCT	1919
QY	1923	GATATTACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTGTTAGCAGAGG	1982
Db	1920	GATATTACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTGTTAGCAGAGG	1979
QY	1983	ACCTGGTATTGCTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTCGTGGGCGCTGGC	2042

Db 1980 ACCTGGTATTCGGTGTGTGTGGAACACAGGGTCTGTCTCAGTGTATCTCTGGGCGCTGGC 2039
Qy 2043 AACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCCTAAAGAACTCTTGACCA 2102
Db 2040 AACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCCTAAAGAACTCTTGACCA 2099
Qy 2103 TGACAGATGTGACACAGACACAGATTTTACAGCTGTACAGCAACACCAATGACTGGCA 2162
Db 2100 TGACAGATGTGACACAGACACAGATTTTACAGCTGTACAGCAACACCAATGACTGGCA 2159
Qy 2163 CTGCTGAATGACCAATGTTCTCCAGGAACCACTGCTGAAAGGCGAGATCTCCAT 2222
Db 2160 CTGCTGAATGACCAATGTTCTCCAGGAACCACTGCTGAAAGGCGAGATCTCCAT 2219
Qy 2223 TTTTAGGTATCAGAAATGCCCCCAAGGATAACCCCTACTACTGTAAACAAGACACAG 2282
Db 2220 TTTTAGGTATGAGAAATGCCCCCAAGGATAACCCCTACTACTGTAAACAAGACACAG 2279
Qy 2283 CTGACAGAGCTGTGCTCTGACACAGAACTGCTGAGTGGAGAGCCCGGAATCAGGAGTGCAT 2342
Db 2280 CTGACAGAGCTGTGCTCTGACACAGAACTGCTGAGTGGAGAGCCCGGAATCAGGAGTGCAT 2339
Qy 2343 TGGCTCTGCCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGG 2387
Db 2340 TGGCTCTGCCGAGTGGCTTGCAGGGTCTCTGTGGTGGTGG 2384

RESULT 14

AAZ91917
ID AAZ91917 standard; cDNA; 2419 BP.

AC AAZ91917;

XX 08-JUN-2000 (first entry)

XX Murine mahogany protein coding sequence akml003.

DE Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
XX weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.

XX Mus sp.

XX WO200005373-A2.

XX 03-FEB-2000.

XX 21-JUL-1999; 99WO-US016484.

XX 21-JUL-1998; 98US-0093630P.

PR 20-OCT-1998; 98US-0104978P.

PR 05-FEB-1999; 99US-00245041.

XX (WILL-) MILLENIUM PHARM INC.

XX Moore K, Nagle DL;

XX WPI; 2000-195103/17.

DR P-PSDB; AAY81804.

XX New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.

XX Claim 1; Fig 8a; 188pp; English.

XX This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in cell-

CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX

SQ Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;

Query Match

Best Local Similarity 39.8%; Score 1521.8; DB 3; Length 2419;

Matches 1899; Conservative 0; Mismatches 247; Indels 205; Gaps 2;

Qy 3 GGTGGCGGACGCGCGGCAACTGAGCAAGGCTGAGGAGGAGGACGGCGGCAACGGCAGC 62
Db 82 GGTGGCGGCGCGCGGCGGCTGAGCGCGGCTGAGGGGGAGCACAGAGACGACAGCAGC 141
Qy 63 GCTCGCGGCGAGGAGCGGCGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
Db 142 GCTCGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 201
Qy 123 GGGGCTGGGGCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 182
Db 202 GGGACCGCGCGCGCGGCTGTCTCGCGCGGCTGTCTCGCGCGGCTGTCTCGCGCGG 261
Qy 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
Db 262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
Qy 243 CGAGGCGCGGCGGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 302
Db 313 CGAGGCGCGGCTGCGGCTGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 372
Qy 303 TGACCGCGCGCTGTCTCAACGGCGGCTGCTGCAACCTGCGACCGCGCGAGTGGCTGCC 362
Db 373 TGACCGCGCGCTGTCTCAACGGCGGCTGCGCAACCTGCGACCGCGCGAGTGGCTGCC 432
Qy 363 CGCGGCTGCGGCGGAGCAATGCGCAGCACTGCGGGGGCGCGCTTCAG----- 410
Db 433 CACGGCTGGGTGGGCGAGCAATGCGCAGCACTGCGGGGGCGCGCTTCAGGACATCTGTCTC 492
Qy 411 ----- 410

Db 493 ACGCTATAATCACAGCTGTTCCGAAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 552
Qy 411 ----- 410
Db 553 GGCTACAGAATAAGTTCAAGATAACCTGGGGCAACTTGGGCTTGTCTCCAAAACCAAAA 612
Qy 411 ----- 410
Db 613 TGAGCGAAAAGAGGCAAGCTAGAGTCTTTTGGGAAAATTTTGGTGAATATTTTTCACC 672
Qy 411 ----ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAATACAAA 466
Db 673 GAGAACTAACTGGCTCTTCTGGATTGTAACAGATGGACCTGGGAAATTTATAATATAGA 732
Qy 467 CGAAGTGCAGCTGGCTCATTGAAGGACAGCCAAATAGAAATATGAGCTTGGTTCAATC 526
Db 733 CGAAGTGCAGCTGGCTCATTGAAGGACAGCCAAATAGAAATATGAGCTTGGTTCAACC 792
Qy 527 ATTTTCTACAGAGTGTAGTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATG 586
Db 793 ATTTTCTACAGAGTGTAGTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTCAC 852
Qy 587 CACCGCTAGTGTCTGCAATTTAGTGGCTCATTTGTTCTCGAGAGATGCGAATGAGACTG 646

Sat Mar 6 08:28:40 2004

us-09-787-097-19.rng

Page 32

Job time : 917.913 secs

Result	Query	Match	Length	DB	ID	Description
1	3713.8	97.2	4072	9	US-09-893-238-16	Sequence 16, Appl
2	3696.2	96.8	8889	9	US-09-893-238-14	Sequence 14, Appl
3	2933.8	77.0	8827	9	US-09-893-238-1	Sequence 1, Appl
4	2245.2	58.9	2625	9	US-09-893-238-18	Sequence 18, Appl
5	1521.8	39.8	2419	9	US-09-893-238-8	Sequence 8, Appl
6	1364.4	35.7	6370	9	US-09-893-238-12	Sequence 12, Appl
7	1225	32.1	6733	16	US-10-197-824-1	Sequence 1, Appl
8	640.2	16.8	1051	9	US-09-893-238-10	Sequence 10, Appl
9	277.4	7.3	625	14	US-10-198-848-5869	Sequence 5869, Ap
10	249.2	6.5	531	9	US-09-864-761-15900	Sequence 15900, A
11	246	6.4	246	9	US-09-736-457-966	Sequence 966, App
12	246	6.4	246	9	US-09-902-941-966	Sequence 966, App
13	246	6.4	246	9	US-09-849-626-966	Sequence 966, App
14	246	6.4	246	14	US-10-017-754-966	Sequence 966, App
15	246	6.4	246	14	US-10-113-871-966	Sequence 966, App

QY
63 GCTCGGGCAGGAGCGCGGCCGACTGGGACTGGGACGTGACCAGGCCTGGAGGCTT

ALIGNMENTS

Db	2280	CTCAGAGAGCTGTGCCCTTGGACACAGAACTCCGACGTGGGAGCCCCGGAAATCAGGAGTGCAT	2339
Qy	2343	TGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAAACTCATGTTTGA	2402
Db	2340	TGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAAACTCATGTTTGA	2399
Qy	2403	AATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAAACCAAAATGC	2462
Db	2400	AATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAAACCAAAATGC	2459
Qy	2463	CCTTTGGCTTCTCTTAACACCCAGAAAGGTAGAAATTTGCTTTAAGCAGCTCGAAT	2522
Db	2460	CCTTTGGCTTCTCTTAACACCCAGAAAGGTAGAAATTTGCTTTAAGCAGCTCGAAT	2519
Qy	2523	AATGCAGTCAATCAGAGCATGTCCAAAGTCTCACTTAAACCCCATTTGGCTCGGCTTCGGAA	2582
Db	2520	AATGCAGTCAATCAGAGCATGTCCAAAGTCTCACTTAAACCCCATTTGGCTCGGCTTCGGAA	2579
Qy	2583	GATCAATGTGTCTTACTGCTGGTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACA	2642
Db	2580	GATCAATGTGTCTTACTGCTGGTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACA	2639
Qy	2643	GTGATGCCGTCTGAGCCCAAGTGAATGCTGTGGAAATTTATCAAGAACCCAGTAC	2702
Db	2640	GTGATGCCGTCTGAGCCCAAGTGAATGCTGTGGAAATTTATCAAGAACCCAGTAC	2699
Qy	2703	TCGGGACTGAAGCTCGAACCTTGCTCAACCCACTCAATGTGTAGTGTCTGTGAAAGGCC	2762
Db	2700	TCGGGACTGAAGCTCGAACCTTGCTCAACCCACTCAATGTGTAGTGTCTGTGAAAGGCC	2759
Qy	2763	TGCAAAACACAGTCTTAAGCAGTCCCGGACACCAATGTGCTTGAAGACAGCATGTGGAGA	2822
Db	2760	TGCAAAACACAGTCTTAAGCAGTCCCGGACACCAATGTGCTTGAAGACAGCATGTGGAGA	2819
Qy	2823	TTGCAACAGCGGACGCTCTGAGTGCATGTGGTGCAGCAACATGAACAGGTGTGACATC	2882
Db	2820	TTGCAACAGCGGACGCTCTGAGTGCATGTGGTGCAGCAACATGAACAGGTGTGACATC	2879
Qy	2883	CAATGCCCTATGTGGCCTCCCTTCCCTTTGGCCAGTGTATGGAATGTTATACGATGAGCAC	2942
Db	2880	CAATGCCCTATGTGGCCTCCCTTCCCTTTGGCCAGTGTATGGAATGTTATACGATGAGCAC	2939
Qy	2943	CTGCCCCCTGAAATTTGTTCAGGCTACTGTACTGTAGTCAATGTCTTGAGCAACGAGG	3002
Db	2940	CTGCCCCCTGAAATTTGTTCAGGCTACTGTACTGTAGTCAATGTCTTGAGCAACGAGG	2999
Qy	3003	CTGTGGCTGTGTACTGATCCGCAATACTTGGCAAAAGGAAATGCAATAGAGGTTTCCCTA	3062
Db	3000	CTGTGGCTGTGTACTGATCCGCAATACTTGGCAAAAGGAAATGCAATAGAGGTTTCCCTA	3059
Qy	3063	TAAAGGACCACTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCAAGCCCT	3122
Db	3060	TAAAGGACCACTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCAAGCCCT	3119
Qy	3123	GCTCAATTTCCAGCATGTCTCTAGAGGACAGCAGATACAATGTGTCTTTCAATCACTGTCC	3182
Db	3120	GCTCAATTTCCAGCATGTCTCTAGAGGACAGCAGATACAATGTGTCTTTCAATCACTGTCC	3179
Qy	3183	AGCTTTGCCAATGCCAACGCCCAAGTAAATGTCATCAATCAGAGCATCTGTGAGAGTGTGA	3242
Db	3180	AGCTTTGCCAATGCCAACGCCCAAGTAAATGTCATCAATCAGAGCATCTGTGAGAGTGTGA	3239
Qy	3243	GAACTTGACCAAGGACGACCTGCGAGACCTGTCATATCTGGCTTCTTACCGTGTATCCAC	3302
Db	3240	GAACTTGACCAAGGACGACCTGCGAGACCTGTCATATCTGGCTTCTTACCGTGTATCCAC	3299
Qy	3303	CAATGAGGGAAATGTCTAGCCCATGCAAGTGAATGGGCAACGCTCTCTGTGCAACACCAA	3362
Db	3300	CAATGAGGGAAATGTCTAGCCCATGCAAGTGAATGGGCAACGCTCTCTGTGCAACACCAA	3359
Qy	3363	CACGGGCAAGTGTCTTGCAACCAACAGGGCGTCAAGGGGACAGGTGCCAGTATGTGA	3422
Db	3360	CACGGGCAAGTGTCTTGCAACCAACAGGGCGTCAAGGGGACAGGTGCCAGTATGTGA	3419

Qy	3423	GGTAGAAATCGATACCAAGAAACCTCTCAGAGAAACATGTATTATACTCTCTTAT	3482
Db	3420	GGTAGAAATCGATACCAAGAAACCTCTCAGAGAAACATGTATTATACTCTCTTAT	3479
Qy	3483	TGACTATCAGTTACACCTTTAGTCTATCCAGGAAGATGTCGTATTACACGCTATCAA	3542
Db	3480	TGACTATCAGTTACACCTTTAGTCTATCCAGGAAGATGTCGTATTACACGCTATCAA	3539
Qy	3543	TTTTTGGCTACTCTCGACGAAACAAACACGGGATTTGGACATGTTCAATGCCTCCAA	3602
Db	3540	TTTTTGGCTACTCTCGACGAAACAAACACGGGATTTGGACATGTTCAATGCCTCCAA	3599
Qy	3603	GAATTTCAACCTCAACATCACTGGGGTGCAGTTTCTCAGCTGGAAACCCAGGCTGAGA	3662
Db	3600	GAATTTCAACCTCAACATCACTGGGGTGCAGTTTCTCAGCTGGAAACCCAGGCTGAGA	3659
Qy	3663	AGAGATCGCTGTGTTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAATGA	3722
Db	3660	AGAGATCGCTGTGTTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAATGA	3719
Qy	3723	GAACTTTGATTTTCGGCAACCCACCAATATCACTTTCTTTGTTATGTCAGTAATTTCCAC	3782
Db	3720	GAACTTTGATTTTCGGCAACCCACCAATATCACTTTCTTTGTTATGTCAGTAATTTCCAC	3779
Qy	3783	CTGCCCCATCAAAATTCAGTGGTGAACCTGAACATGA	3819
Db	3780	CTGCCCCATCAAAATTCAGTGGTGAACCTGAACATGA	3816

RESULT 2

```

US-09-893-238-14
Sequence 14, Application US/09893238
Patent No. US20020150973a1
GENERAL INFORMATION:
APPLICANT: Moore, K.
APPLICANT: Nagle, D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
FILE REFERENCE: 7853-237
CURRENT APPLICATION NUMBER: US/09/893,238
CURRENT FILING DATE: 2001-06-27
PRIORITY APPLICATION NUMBER: 09/245,041
PRIORITY FILING DATE: 1999-02-05
PRIORITY APPLICATION NUMBER: 60/093,630
PRIORITY FILING DATE: 1998-07-21
PRIORITY APPLICATION NUMBER: 60/104,978
PRIORITY FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 129
SOFTWARE: FastSeq for Windows version 3.0
SEQ ID NO 14
LENGTH: 8589
TYPE: DNA
ORGANISM: Homo sapiens
US-09-893-238-14

```

	Query Match	96.8%	Score 3696.2	DB 9	Length 8589
	Best Local Similarity	98.5%	Pred. No. 0		
	Matches 3744	Conservative	0	Mismatches 48	Indels 9
	Gaps				1
Qy	3	GGTGGCGCAGCGCGGCAACTGAGGCAAGCGTGAAGAGGAGCGCGGCGACGCAGC	62		
Db	9	GGTGGCGCAGCGCGGCAACTGAGCAAGCGTGAAGAGGAGCGCGGCGACGCAGC	68		
Qy	63	GCTCGCGGCGAGGACGCGCGCGCACTGGGACTGGGACTGACAGGCTGGAGGCC	122		
Db	69	GCTCGCGGCGAGGACGCGCGGCGCGACCGACCCGTGACCGGCGCTGGAGGCC	128		
Qy	123	GGGGCTGGGGCGCGGGCTGCGCTCCCGCGGTGCTGTCTCCACCGCTCGGCGCACGGCT	182		
Db	129	GGGACCGCGCGCGCGCTGTGTCTCCCGGGGTGCTGTGCGGGGCTGCGCCCGCGCC	188		
Qy	183	GCTGCTGCTGTGTGTGTGCTCCCGCGCGCGCTGTGTGCTGTGCTGTGCTGTGAGGC	242		

Db 313 CGAGGCGCTGTGCGGTGGCGCGCGGTGTCCGGCTCGGCGCGAGCGAGGCCAAGGAATG 372
QY TGACCGGCGCTGTGTCAACGGCGGTGCTGCAACCGCTGCGACCGCGCGAGTGGTCTGGCC 362
Db 373 TGACCGGCGGTGTGTCAACGGCGCGGTGCTGCAACCGCTGCGACCGCGCGAGTGGTCTGGCC 432
QY CGCGGCTGGGTGGCGAGCAATGCGAGCACTGCGGGGCGCGCTTCAG----- 410
Db 433 CACGGCTGGGTGGCGAGCAATGCGAGCACTGCGGGGCGCGCTTCAGGACATCTGTCTC 492
QY 411----- 410
Db 493 AGCCCTAATCACAGCTGTTCCGAAGTGAAGGTGGAGGAACAGTTCCAGGCAAGCTTC 552
QY 411----- 410
Db 553 GGCACAGAAATAGTTCAAGAGTAACCTGGGGCACTGGGGCTTGTCTCCAAACCAAAA 612
QY 411----- 410
Db 613 TGAGCGAAAGAGCAAGCTAGAGTCTTTTGGGAAAAATTTAGCTGCACTAATTTTTCACC 672
QY 411-----ACTAACTCGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAACAAAA 466
Db 673 GAGAACTAATCTGCTCTTCTGGATTTGTAAACAGATGGACCTGGGAAATTAATAATAAGA 732
QY 467 CGAAGTGACGCTGGTCAATGGAAGGACAGCCAAATAGAAATAGAGACTTCTGTTTCAATC 526
Db 733 CGAAGTGACATGGCTCAATGGAAGGACAGCCAAATAGAAATAGAGACTTCTGCTTCAACC 792
QY 527 ATTTTGTACAGATGATGCGGACCAATTAATATATATATATATATATATATATATATATAT 586
Db 793 ATTTTGTACAGATGATGCGGACCAATTAATATATATATATATATATATATATATATATAT 852
QY 587 CACCGCTAGTCTGCATTTAGTGGCCTCAATGTTCTCGAGAGATGGCAATGAGACTG 646
Db 853 CACCTCTGATCTGCTGCTTTAGTGGCCTCAATGTTCTCGAAGAGATGGCAATGAGACGG 912
QY 647 TCCCTGAGTTTGTGCGACATCAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
Db 913 CTCCTGAGTCACTGCTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
QY 707 ATAAATGACTGCAATTAAT 766
Db 973 ATAAATGACTGCAATTAAT 1032
QY 767 GCGGAGGAGGTGAAGATAGTAATAGCAGCGAAACTGTTGAATGTAATGTTCTGAAA 826
Db 1033 GCGGAGGAGGTGAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAA 1092
QY 827 ACTGAAAGGTGAACATGTGACATCTCTCACTGTACAGCAACTGTTGTTTCTCTCATC 886
Db 1093 ACTGAAAGGAGGTGCTGTGACATCTCTCACTGTACAGCAACTGTTGTTTCTCTCATC 1152
QY 887 GAGGCATCTGCAATTCAGATGTACAGAGATGCTCTGCTTCTCAGACTGGCAGGGTC 946
Db 1153 GAGGCATCTGTAATGCAAGCATACAGAGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1212
QY 947 CTGAGTGTTCAGTCTCTGACAGTCACTGCTGTACAGCAACTGTTGTTTCTCTCATC 1006
Db 1213 CTGAGTGTTCAGTCTCTGACAGTCACTGCTGTACAGCAACTGTTGTTTCTCTCATC 1272
QY 1007 ACTTAAAGCTCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACCAATTAATGTTGGTTG 1066
Db 1273 ATTTAAAGCTCCAGAGCTCTCATAAAGCTGTGGTCAATGGAACCAATTAATGTTGGTTG 1332
QY 1067 TTGAGGATATATGTTCAACACTCAGATTAATACATGTTTCTAGCGTATGACCTTGCCTT 1126
Db 1333 TTGCGGATATATGTTCAACCACTCAGATTAATACATGTTTCTAGCGTATGACCTTGCCTT 1392
QY 1127 CTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACCAATGTTGTTTGTAGATATGTTGCTATT 1186
Db 1393 CTAGGAGTGGCTTCCACTAAACCACTTCTGTGAACCAATGTTGTTTGTAGATATGTTGCTATT 1452

QY 1187 CTTTGGCATTTATACAGGATAAAATTTACATGTATGGAGGAAAAATTTGATCCAACTGGGA 1246
Db 1453 CTTTGGCATTTATACATAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTTCAACAGGGA 1512
QY 1247 ATGTGACCAATGAGTTGAGAGTTTTCACATTTCAATATGATCATGGGTGTTGTTGACCC 1306
Db 1513 ACGTGACCAATGAGTTGAGAGTTTTCATATTTCAATATGATCATGGGTATTTGTTAACTC 1572
QY 1307 CTAAGGCAAAAGGAGCAGTATGAGTGGTGGGCACTCTGACACATTTGTTACTGAAGA 1366
Db 1573 CGAAAGTAAGGATCAGTATGAGTGGTGGGCACTCAGCACACATTTGTTACTGGCAT 1632
QY 1367 ATGCGCGAGTGGTCACTGCTGCTCATCTTTGGTCACTGCGCTCTCTATGATATATAAGA 1426
Db 1633 CTGGCGTGGTCACTGCTGCTCATCTTTGGTCACTTGGTCACTCTATGATATATAAGG 1692
QY 1427 ATGTGACGGAATATGATTTGGATAAGAACACATGAGAGTATATTACACCCAGGGTGGCC 1486
Db 1693 TTGTGACGGAATATGATTTGGAAAGAACACATGAGAGTATATTACATCTCAGGGTGGTC 1752
QY 1487 TTGTGAAAGGGGTTAGCGCATAGCAGTGTTTACGACCATAGGACCGGCGCTATAG 1546
Db 1753 TTGTGAAAGGGGTTAGCGCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACG 1812
QY 1547 TTCATGGTGGCTACAAAGGCTTTTCAGTCCCAATAAGTACCAGGCTTGAGATGATCTCTACC 1606
Db 1813 TTCATGGTGGCTACAAAGGCTTTTCAGCGCCAAACAATAATACCGCTTGAGATGATCTCTACA 1872
QY 1607 GATATGATGTGATACCCAGATGTGACCATCTTCTTAAGACACGCGATTTTTCGGTTACT 1666
Db 1873 GATACATGTGATACCCAGATGTGACCATCTTCTTAAGACACGCGATTTTTCGGTTACT 1932
QY 1667 TGACACAGCTGTGATGAGTGAAGCAACCATGCTGTTGTTGGGGAACACACACAATG 1726
Db 1933 TGACATACAGCTGTGATGAGTGAAGCAACCATGCTGTTGTTGGAGGAACACACACAATG 1992
QY 1727 ACACATCTATGAGCCATGCGGCAAAATGCTTCTTTAGATTTTATGAGCTATGACATTTG 1786
Db 1993 ACACCTTCCATGAGCCACGGTGCACAAATGCTTCTCTCGGACTTCATGGCTTATGACATTTG 2052
QY 1787 CCGTGACCGCTGCTGAGTGTGAGTGAAGCAACCATGCTGTTGTTGGGGAACACACACAATG 1846
Db 2053 CTTGTGACCGCTGCTGAGTGTGAGTGAAGCAACCATGCTGTTGTTGGAGGAACACACACAATG 2112
QY 1847 GGCATTCAGCAGTCTTTPACACACAGCACCATGATGTGTGTTGGTGGTTCATATGATCTCTCC 1906
Db 2113 GGCATTCAGCAGTCTTTPACACACAGCACCATGATGTGTGTTGGTGGTTCATATGATCTCTCC 2172
QY 1907 TCCCTGACGACATCTTGGTATTTCCCTCGGACAGTGTGATGCGCATGGAGTGAAGCG 1966
Db 2173 TCCCTGAGTGGTCTTGGTCTTTTCTCGGAGTGGGATGCAACCGCAGTGAAGCTG 2232
QY 1967 CTTGTTTACAGCAGGACCTGCTATTTCCGTGTGTGTGGAACACACAGGCTGCTCTCAGTGT 2026
Db 2233 CTTGTTGTCAGCAGGACCTGCTATTTCCGTGTGTGTGGAACACACAGTGTCTCTGATGTA 2292
QY 2027 TCTCGTGGCGCTGGCACTGATGAAAGAAAGTTAAATCAAGATGTTTTCCTA 2086
Db 2293 CTTCTCGGAGTGGCACTGAAAGAAAGTTAAATCAAGATGTTTTCCTA 2352
QY 2087 AAAGAACTCTTCACCATACAGATGTGACAGCACACAGATTTGTTACAGCTGTACAGCA 2146
Db 2353 AAAGAACTCTTCACCATACAGATGTGACAGCACACAGATTTGTTACAGCTGTACAGCA 2412
QY 2147 ACACCAATGACTGCGACTGCGTGGTGAATGACCATTTGTTCCCGAGGAACACACAGTGTCTCAG 2206
Db 2413 ATACCAATGACTGCGACTGCGTGGTGAATGACCATTTGTTCCCGAGGAACACACAGTGTCTCAG 2472
QY 2207 AAGGCCAGATCTCCATTTTATAGGTATGAAATTTGCCCAAGGATAACCCCATGTACTACT 2266
Db 2473 AAGGCCAGATCTCCATTTTATAGGTATGAAATTTGCCCAAGGATAACCCCATGTACTACT 2532

Db 2340 TGCCCTGCCGCTAGGCTTGCAGGGTCATCTTGGTGTGTGGG 2384

|||||

RESULT 5

US-09-893-238-8

; Sequence 8, Application US/09893238

; Patent No. US20020150973A1

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

; FILE REFERENCE: 7853-237

; CURRENT APPLICATION NUMBER: US/09/893,238

; CURRENT FILING DATE: 2001-06-27

; PRIOR APPLICATION NUMBER: 09/245,041

; PRIOR FILING DATE: 1999-02-05

; PRIOR APPLICATION NUMBER: 60/093,630

; PRIOR FILING DATE: 1998-07-21

; PRIOR APPLICATION NUMBER: 60/104,978

; PRIOR FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8

; LENGTH: 2419

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-893-238-8

Query Match 39.8%; Score 1521.8; DB 9; Length 2419;

Best Local Similarity 80.7%; Fred. No. 0;

Matches 1895; Conservative 0; Mismatches 247; Indels 205; Gaps 2;

Qy 3 GGTGCGCGCAGCGCGGCAACTGAGCGAAGGTGTGAGGAGGAGCGCGGCGAGCGAGC 62

Db 82 GGTGCGCGCGCGCGGCGGCGCTGAGCGCGCTGAGGCGGAGCAGGAGCGAGCAGCAGC 141

Qy 63 GTCGCGGCGAGGAGCGCGGCGCGGCGGCGCTGAGGAGGAGCGCGGCGGCGAGCGC 122

Db 142 GCGTGGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 201

Qy 123 GGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 182

Db 202 GGGAGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 261

Qy 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242

Db 262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312

Qy 243 CGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302

Db 313 CGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 372

Qy 303 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 362

Db 373 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 432

Qy 363 CGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410

Db 433 CACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 492

Qy 411 ----- 410

Db 493 AGCGCTATAATCACAGCTGTCGGAAGGTGAGGCTGGAGAACAGTTCGAGGCAAGCTTC 552

Qy 411 ----- 410

Db 553 GGCTACAGATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTCCAAAACCAAAA 612

Qy 411 ----- 410

Db 613 TGAGCGAAAAGGAGCAGCTAGAGTCTTTTGGGAAAATTTTAGCTGACTAATTTTTCACC 672

Qy 411 ----ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAACAAA 466

Db 673 GAGAACTAACTGGCTCTTCTGGATTGTGTAACAGATGGACCTGGGAATTAATAATAAGA 732

Qy 467 CGAAGTGCAGTGGCTCANTGAAGGACAGCCAAATAGATATGAGACTTCGTTTCAATC 526

Db 733 CGAAGTGCAGTGGCTCANTGAAGGACAGCCAAATAGATATGAGACTTCGTTTCAATC 792

Qy 527 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATG 596

Db 793 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGAGCTCAATCTACG 852

Qy 587 CACCGTGTAGTGTGCTTGTAGTGGCTCANTGTTCTCTGAGAGAGATGCAATGAGACTG 646

Db 853 CACCTCTGATTGTGCTTGTAGTGGCTCANTGTTCTCTGAAAGAGATGCAATGAGACGG 912

Qy 647 TCCTGTGAGGTTGTGCGACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTT 706

Db 913 CTCTGTGAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 972

Qy 707 ATAATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766

Db 973 ATAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032

Qy 767 GCGGAGGAGTGTAGATCAGTAAATAGCAGCGGAACTGTTGAATGTTGAATGTTCTGAA 826

Db 1033 GCGGAGGAGTGTAGATCAGTAAATAGCAGCGGAACTGTTGAATGTTGAATGTTCTGAA 1092

Qy 827 ACTGGAAGGTGAAGCATGTGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886

Db 1093 ACTGGAAGGTGAAGCATGTGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152

Qy 887 GAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946

Db 1153 GAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1212

Qy 947 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1006

Db 1213 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272

Qy 1007 ACTTAAAGCTCCCGAGAGCATCTAATAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066

Db 1273 ATTAAAGCTCCCGAGAGCATCTAATAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332

Qy 1067 TTGAGATATATGCTTCAACCATCTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126

Db 1333 TTGCGGATATATGCTTCAACCATCTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392

Qy 1127 CTAGGAGTGGCTTCCACTAAACCGTTCTGTAACATGCTGCTGCTGCTGCTGCTGCTGCT 1186

Db 1393 CTAGGAGTGGCTTCCACTAAACCGTTCTGTAACATGCTGCTGCTGCTGCTGCTGCTGCT 1452

Qy 1187 CTTTGGATATATCAAGGATAAATTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1246

Db 1453 CTTTGGATATATCAAGGATAAATTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1512

Qy 1247 ATGTGACCAATGAGTGTGAGAGTGTGCTGCAATTCATATGATGCTGCTGCTGCTGCTGCT 1306

Db 1513 ACGTGACCAATGAGTGTGAGAGTGTGCTGCAATTCATATGATGCTGCTGCTGCTGCTGCT 1572

Qy 1307 CTAAGGCAAGGAGGAGGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366

Db 1573 CGAAAGCTAAGGATCAGTGTGAGAGTGTGCTGCAATTCATATGATGCTGCTGCTGCTGCTGCT 1632

Qy 1367 ATGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426

Db 1633 CTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692

Qy 1427 ATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1486

Db 1693 TTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1752

QY 3254 CAGCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACAAATGGAGGA 3313
Db 1021 CAGCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACAAATGGAGGA 1080
QY 3314 AATGTACGCCATGCAAGTGCATGGGACCGGTCTCTGTGGACACCAACACCGGGCAAGT 3373
Db 1081 AATGTACGCCATGCAAGTGCATGGGACCGGTCTCTGTGGACACCAACACCGGGCAAGT 1140
QY 3374 GCTTCTGCACCAACCAAGGGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAATC 3433
Db 1141 GCTTCTGCACCAACCAAGGGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAATC 1200
QY 3434 GATACCAAGGAACCTCTCAGAGGAACATGTTATTAATCTCTCTTATTGACATACAGT 3493
Db 1201 GATACCAAGGAACCTCTCAGAGGAACATGTTATTAATCTCTCTTATTGACATACAGT 1260
QY 3494 TCACCTTTAGTCTATCCAGAGGATGATCGCTATTAACAGCTATCAATTTTGGGCTA 3553
Db 1261 TCACCTTTAGTCTATCCAGAGGATGATCGCTATTAACAGCTATCAATTTTGGGCTA 1320
QY 3554 CTCTCTGACGAACAAACAGGGATTGGACATGTTTCATCAATGCTCCCAAGAAATTCACCC 3613
Db 1321 CTCTCTGACGAACAAACAGGGATTGGACATGTTTCATCAATGCTCCCAAGAAATTCACCC 1380
QY 3614 TCACATCAGCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGAGAGATGCTG 3673
Db 1381 TCACATCAGCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGAGAGATGCTG 1440
QY 3674 TTGTTTCAAAACCAACAAATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATT 3733
Db 1441 TTGTTTCAAAACCAACAAATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATT 1500
QY 3734 TTGCAACACCCCAATATCACTTTCTTTGTTTATGTCAAGTAATTCACCTGGCCCATCA 3793
Db 1501 TTGCAACACCCCAATATCACTTTCTTTGTTTATGTCAAGTAATTCACCTGGCCCATCA 1560
QY 3794 AAATTCAGT 3803
Db 1561 AAATTCAGT 1570

RESULT 7

US-10-197-824-1
; Sequence 1, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
; FILE REFERENCE: 9U 106 R1
; CURRENT APPLICATION NUMBER: US/10/197,824
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (575)..(4711)
; OTHER INFORMATION:
US-10-197-824-1

Query Match 32.1%; Score 1225; DB 16; Length 6733;
Best Local Similarity 61.9%; Pred. No. 0;
Matches 2175; Conservative 0; Mismatches 1250; Indels 90; Gaps 11;
QY 332 GCAACCTGGCCAGCCAGTGGCTCTGCCCGCCGGCTGGGTGGGGAGCAATGCCAGC 391
Db 789 GCGCTGTGTCACTCCACCTGCCCTCGGACCCGGCTGGGTGGGGAGCCAGTCCAGC 848
QY 392 ACTGGGGGGCGCGCTTCAGACTAACTGGATCTCTCTGGGTTTGTGACAGATGGACCTGGAA 451

Db 849 ACTGCCAGGGCAGGTTCAAGTTAAAGAACCTTCTGGATATTTAAAGATGCCCAATTA 908
QY 452 ATTATAAATACAAACGAAAGTGCAGCTGCCTCATTCAGGACAGCAAAATAGATAATGA 511
Db 909 ACTATAAATATAAATAAATTAATCTACTTGGCTCATTCAGAGGCTATCAAAATCAGTGTAA 968
QY 512 GACTTCGTTTCAATCATATTTTGTCTACAGAGTGTAGTTGGACCAATTTATATGTTTATG 571
Db 969 GATTAAAGATTCATCATATTTTGTCTACAGAAATGTAGCTGGGATCATATGTATGTTTATG 1028
QY 572 GGGACTCAATTTATGCAACCGCTAGTTGTCTGCTATTTAGTGGCTCATTTGTCTCTGAGAG 631
Db 1029 GAGATTCAATATAATGACACCTTTAATAGCTGTACTTAGTGGTTGATAGTCCCTGAAATA 1088
QY 632 ATGGCAATGAGACTGTCCCTGAGGTTGTGCAACATCAGGTTATGAGCTTCTGCTGCAATTTT 691
Db 1089 GGGCAATGAAACTGTGCTGAAGTTGTACTACATCTGGCTATGCACTGTGTACATTTT 1148
QY 692 TTAGTCAATGCTGCTTATATTTTGTCTGATTTTAAATTAATTAATTAATTAATTAATGTC 751
Db 1149 TTAGTCAATGCTGCTGCTTATATCTAATCTGTTTCAACATTTTCTATTAATCAATCTTGT 1208
QY 752 CAAATAACTGCTCAGGCCGAGGAGAGTGTAAAGATCAATAGTAAATAGCAGCGAAACT---GTTG 808
Db 1209 CTAACAATTTGCTCTGCTCATGGAGAGTGTACAACATAGTGTCTGTCTCCAAGTCAAGTAT 1268
QY 809 AATGTGAATGTTCTGAAAACCTGAAAGGTGAAGCATGTGACATTCCTCTCTCTGACAGACA 868
Db 1269 AATGTGAATGTGATAAATACTGGAAGGTGAAGCTTGTGATATTCCTCTACTGTAAAGCCA 1328
QY 869 ACTGTGCTTTTCTCTCATCGAGGCATCTGCAATTCAGATGATGTGAGAGATGCTCCTGCT 928
Db 1329 ATTGGCGCAGTCCAGATCAGGTTACTGTGACCTGATGGAGAAAATTTATGTCTGCA 1388
QY 929 TCTCAGATGGCGAGGTCTCTGAGTGTTCAGTTCCTGTACAGCTAACAGTCACTTTTGGGA 988
Db 1389 ATGATAGTTGGCAAGGTCTCTGATTTCTTTGAATGTCCCTCTACTGAGTCTTACTGGA 1448
QY 989 CTCGAGAGGATATTTCTAATTTAAAGC-----TCCCCAGAGCATCTCATAAAGCTGTGG 1042
Db 1449 TTTCTGCCAAACGTTTAAACCTTCTAGTCTCTCTAGTGGGCTTCAATAAAGCAGTTT 1508
QY 1043 TCAATGGAACATTTATGTGGTTTGTGGAGGATATATGTTCAACCACTCAGATATAACA 1102
Db 1509 TACACGGGAATTTATGTGGTGCATTTGGTGGATATATCTTTAACTACAGTCTCTTTCAAA 1568
QY 1103 TGGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCGTTCCTGTGAACA 1162
Db 1569 TGGTCTCTAAATTTACAATTTAGAAAGCAGTATATGGAATGTAGGAATCCATCAAGGGGAC 1628
QY 1163 ATGTGGTTGTAGATATGGTCAATTTTGGCATTTATCAAGGATAAAATTTACATGTATG 1222
Db 1629 CT---CTCCAGAGATATGGACACTCTCTGCTTTTATATCAGGAAAACATCTTTATGTATG 1685
QY 1223 GAGGAAAAATTTGATCCAA---CTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTC 1279
Db 1686 GAGGAGAAATTTGAAAACAAATGATGGCAATGTCAAGATGAATTTATGGTTTAAACATAC 1745
QY 1280 ATATGAGTCAATGGGTGTCTGACCCCTAAGGCAAAAG-----AGCAGTATGTCAG 1330
Db 1746 ATAGTCAATGAGGATGACAAATCTCTACTCTTCTTGGACATGGTCAAGTATGCTG 1805
QY 1331 TGGTTGGGCACTCTGCAACATTTGTACACTGAAGAATGCCGAGTGGTCAATGCTGGTCA 1390
Db 1806 TGGAGGGACATTCAGACACATATTTATGGAGTTGGATAGTAGATGTTGTCTCATGATCAAA 1865
QY 1391 TCTTTGTCTACTGCCCTCTCTATGGATATATAGCAATGTGCGAGGATATGTTTGGATA-1450
Db 1866 TATTTGGATATCTCGCAATATATGTTTATACAGAGCATACAGGAATACCATATCTCAT 1925
QY 1451 AGAACACATGAGTATATTTACACACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATA 1510

Db 1926 CAAACACTTGGCTGTTCCAGAAAATAAAGAGAGCTATTGTACAAGGTGATATGCCAATA 1985
Qy 1511 GCAGTGTTTACGACCAATAGGACAGAGCGCCCTATACGTTTCATGTTGCTTACAGGCTTTCA 1570
Db 1986 CTAAGTGTATGATGAATAAACAAGTCCATTTATGTTTCATGAGGGTATAAAGCAATGTC 2045
Qy 1571 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGATACCCAGATGT 1630
Db 2046 CAGGGAAACAAATATGGAATGGTGTGATGATCTTTATAAATATGAAGTTAACTAAGACTT 2105
Qy 1631 GGAACATCTTAAGGACACCGCAATTTTCCGTTACTTGCACACAGCTGTGATGTAGTGT 1690
Db 2106 GGAATTTTGAAGAAGTGGTGTGCGAGATACCTTCAATTCAGCTGTTCTTATCAATG 2165
Qy 1691 GAAACATGCTGGTGTGAGGGAACACACACAATGACACATCTATAGAGCAATGCGGCA 1750
Db 2166 GAGCTATGCTTATTTTGGAGGAAATACCCATAATGACACTTCTTGTAGTAACGGTGCAA 2225
Qy 1751 AATGCTTCTTTCAGATTTTCAATGGCTATGACATTCCTGTGACCGCTGTGAGTCTTC 1810
Db 2226 AATGTTTTTCTGCCGATTTCTGGCATATGACATAGCTTGTGATGAATGGAATAATCTAC 2285
Qy 1811 CAGACCTGATCTCCACCAATGATGACACAGATTTGGCCATTCAGCAGTCTTTACACAACA 1870
Db 2286 CAAAACCAATCTTCATAGAGATGTCAACAGATTTGGACACTCTGCACTAGTCAATTAACG 2345
Qy 1871 GCACCATGATGTGTCGTTGTTTCAATAGTCTCTCTCAGCAGACATCTCTGATTTCA 1930
Db 2346 GGTCCATGATATATTTTGGGGATTTTCTAGTGTACTCTTAATGATATCTTGTATACA 2405
Qy 1931 COTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTGTAGCAGCAGACTGTGTA 1990
Db 2406 AGCTCCCAATTTGCAAGGCTTTCACAGATGAAGACTTTGTAAAAATGCTGCTCCAGGGA 2465
Qy 1991 TTCGTGTGTGGAACACAGGCTGCTCTAGTGTATCTCGTGGCGCTGCACTAGTATG 2050
Db 2466 TAAATGTGTTTGGAAATAAATACTGTGAAT-----CTTGGGAATCTGGGA 2513
Qy 2051 AACAAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTTGCACCATCAGAGAT 2110
Db 2514 ATACTAAATAATCTTAGAGCAAGTGCCTCTTAAACAGCTGCTTCTGATGACAGAT 2573
Qy 2111 GTGACGACACAGATTTTACAGCTGTACAGCAACACCAATGACTGCCACTGGTGCA 2170
Db 2574 GTTACAGATATGCAATGTGTCAGCTGTACTGTGCAATACAAATGGGTGCCAATGGTGTG 2633
Qy 2171 ATGACCAATGTGTCGCCAGAACACAGCTGCTCAGAGGGCCAGATCTCCATTTTAGGT 2230
Db 2634 ATGACAAGAAATGCAATTCGGCAATATGTAATCTGAGTGTCTGTCAAGAACTACACCA 2693
Qy 2231 ATGAGAAATGGCCCAAGGATAACCCATGTACTACTGTAAACAAGAACAGCTGCAGGA 2290
Db 2694 AATGTCAATGTGAGAAATCA-----GCAGATTTGTAAACAATTTACAGCTGTAAA 2744
Qy 2291 GCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGCAATCAGAGTGCATTTGCCCTGTC 2350
Db 2745 GCTGTTCCTAACTTGAATTTGCCAGTGGGATCAGACAGCAAGAATGCCAGCTTTAC 2804
Qy 2351 CCGAAATATCTGTGGCAATTTGCTGGCAATTTGTTGGAAATCTCATGTTTGAATAATTA 2410
Db 2805 CAGCTCATCTTTGTGGAGAGATGGAGTCAATTTGGGATGCTTGTCTAGAGTCAAT 2864
Qy 2411 CTGCCAAGAGAAATATGACAAATGTAATTTGTTCTGTAGGAACCAACAATGCCCTTTTGG 2470
Db 2865 CCAGTAGAGAAATCTATGACAAATGCAAACTTTATGCTATAATCTTTAGTGGAAATCTTG 2924
Qy 2471 CTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTCTTAAGCAGCTGCGAATAATGCA 2530
Db 2925 CTTCAATTAACAACTTCAAAAGAGTAGAATTTGTT-----TGAGTG 2966
Qy 2531 CATCTCAGAGCATGTCCAAAGTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 2590
Db 2967 AAATACAGAAATATACAAACAGAAAGTATCACCTTGGTAGGTTCGGCAAGATCAATA 3026

Qy 2591 TGTCTACTGCTGCTGGAAAGATATGTCCCAATTTTACAAATAGTTTTTACTACAGTGATGC 2650
Db 3027 TATCTATTGGGATGGAAAGACATGCTCTCTTTTACAAACAACACTACAGTGGCTTC 3086
Qy 2651 CGTCTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGAC 2710
Db 3087 CTGCGGAACCAATGATTTCTGGGTTTGTGCAATCTGGAAGGGCTGAGTGGCAGGT 3146
Qy 2711 TGAAGGTGCAACCTGTCATCAACCCCACTCAATGGTAGTGTCTGTGAAAGCCCTG----- 2764
Db 3147 TAAAGCTAATCTTGTACATCTATGCAAAATGCGCTTGTCTGTGAAAACCTGTTGTTA 3206
Qy 2765 ---CAAACACAGTGTCTGAGTGTGCGGACACATGTCCTTGTAGCAACAGCATGTTGGAG 2821
Db 3207 GTCCAAATCAAAATGGGAGCGCTGCAAAAGCAATGCTCTCTGAGGACATCATGTTCCA 3266
Qy 2822 ATTGCACAGCGGCTGCTCTGAGTGTGCTGAGTGTGAGCAACATGAAGCAGTGTGTGACT 2881
Db 3267 ACTGTACAAGCAATGCAATGGAGTGTATGTGTGAGCAGTACGAAACGATGTTGACT 3326
Qy 2882 CCAATGCTATGTGGCTCTCTCTCTTTTGGCCAGTGTATGGAATGTTATACATGAGCA 2941
Db 3327 CTAATGCTATATCATCTCTTTTCCATATGGAATGCTAGAGTGGCAAC---TGCCA 3383
Qy 2942 CCTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCTATGCTTGGAGCAACCA 3001
Db 3384 CCGTCTCTCTCTTAAATTTGTTGGAATGGAACCTGTGAGCAGTGTGGAACAGCTG 3443
Qy 3002 GCTGTGGCTGGTACTGCCAGCAATCTGGAAGGGAAATGATAGAGGTTCT 3061
Db 3444 GATGTGGCTGGTCAATGATCTTAGTAATACAGGAAGAGACATGCAATTTGAAGTTCT 3503
Qy 3062 ATAAAGACAGTGAAGATGCTTCCAGAGCCCTACAGGAATTTCTATCCACAGCCCC 3121
Db 3504 CACGGGACCAATGAAG-----CTTATTGGAATGCCAACATGATGG 3548
Qy 3122 TGTCTAATTTCCAGCATGCTGTAGAGCAGCAGATACAACTGGTCTTTTCAATCTCATGTC 3181
Db 3549 TTTCTGACACCAATCTTTCGCCCAAGAAAGAACTATAGTGGTCTTTATCCAGTGT 3608
Qy 3182 CAGCTTGCAATGCAACAGGCAAGTAATGCAATCAATCAGAGCATCTGTGAAAGTGTG 3241
Db 3609 CAGCTTGCCAGTGTATGACATAGCACTTGCAATCAATTAATGTGCGAACAAGTGT 3668
Qy 3242 AGAACTGACACAGCAAGCACTGAGACCTGCAATATCTGGCTTCTACGCTGATCCCA 3301
Db 3669 AATCTCTACCAAGAAAGCAGTGTCAAGTTGTATGCCAGTTTATTTAGAGATCCAA 3728
Qy 3302 CCAATGAGGGAAATGTCAGCCATGCAATGGCAATGGCAACCGCTCTCTGTGCAACCA 3361
Db 3729 CCAATGGTGACAGTGCACAGCTTGTACATGCACTGAGTGGCCATGCAAAATATCTGCTG 3788
Qy 3362 ACAGGGCAAGTCTTCTGCACCAACAGGCGCTCAGGGGACAGTGGCCAGCTATGTG 3421
Db 3789 ACACAGAAAAATGTTTCTGCACAACTAAAGGAATAAAGGTGACCAATGCCAATTTATG 3848
Qy 3422 AGGTAGAAAATCGATACCAAGAAACCCCTCTCAGAGGAACATGTTTATTAATCTTCTTA 3481
Db 3849 ACTCTGAAAATCGCTATGTTGTAATCCACTTAGAGGAACATGTTTATACAGCCTTTGA 3908
Qy 3482 TTGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATCGCTTATTTACAGCTATCA 3541
Db 3909 TTGATATCAATTTTACCTTTAGCTTATTTACAGGAAGATGATGCCACCATACTGCCATA 3968
Qy 3542 ATTTTGTGCTACTCTCTGACGAACAAACAGGCAATTTGCAATGTTTCAATGCTCTCA 3601
Db 3969 ACTTTATAGCAACCCAGCAACAGTGCACAAAATCTGGATATATCAATTAATGCAATCA 4028
Qy 3602 AGAATTTCACTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGAG 3661
Db 4029 ACAACTTTAATCTCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4088

3662	AGACATGCGTGTGTTTTCAAAACCAACATTAAGAGTACAAAGATAGTTCTCTAATG	3721
QY		
4089	AGAGACCTCTATAGTTTCCAGAAATATATAAAGAANTACAGAGATAGTTTTCCTATG	4148
Db		
3722	AGAATTGTGATTTTGGCAACCAACCCCAATATCACTTTCTTTGTTATGTCAGTAATTTCA	3781
QY		
4149	AAAAATTTAACTTTTAGAAGCAATCCTAACATTACATTTCTATGTGTCGTGACCAACTTT	4208
Db		
3782	CCTGGCCCATCAAAATTTCAAGTGCAGTAACGAAACA	3816
QY		
4209	CCTGGCCTTATAAATACAGATGTCATTTCTCAAA	4243
Db		

```

RESULT 8
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-10

```

[illegible]

D	b		627	TTC	TGGATTGTTAA	CAGATGGACCTGGGAATTATAAATATAAGACGAGTGCACATGGCT	686	
Q	y		483	CAIT	GAGGACAGCCAAATAGAATAATGAGACTTCGTTC	CAATCATTTTGCTACAGAGTG	542	
D	b		687	CAIT	GAGGACAGCCAAATAGAATAATGAGACTTCGCTTCA	CCATTTTGCTACAGAA TG	746	
Q	y		543	TAG	TGGGACGANTTATATGTTTATGATGGGACTCAATTTATG	CACCCTAGTTGCTGC	602	
D	b		747	TAG	CTGGGACCAATTATATGTTTTATGATGGGACTCAATCTAC	GCACCTCTGATTTGCTGC	806	
Q	y		603	ATT	TAGTGGCTCATTCTCCTGAGAGAGATGCCAATGAGACTGTCCCT	CAGGTTGTTGC	662	
D	b		807	CTTT	TAGTGGCTCA TTGTTCTGTAAGAGATGGC	NATGACCGCTCTGAGGTCACTGT	866	
Q	y		663	CACAT	CAGGTTATGCTTGTGTCATTTTTTTTAGTGATGCTGCTTATAA	TTTGA CTGGATT	722	
D	b		867	CAC	TT	CAGGTTATGCACTGCTGCA TTTTTCAGTGATGCTGCTTATAA	TTCTGACTGGATT	926
Q	y		723	TAAT	ATTACTTTCAGTTTTTGATATGTTGTCCAATAACTGCTCAGG	CCGAGGAGAGTGTA A	782	
D	b		927	TAAT	ATCACCTTACAATTTTGACATGTGTCCGAATAATTTGCTCAGG	CCGAGGAGAGTGTA A	986	
Q	y		783	GAT	CAGTAATAGCAGCGAAAAC TGTGTAATGTGTAATGTTCTG	AAAACTGGA AAAGGTGAAGC	842	
D	b		987	GAG	CAGTAACAGCAGACCGCTGTTGAGTGTGTAATGTTCTG	AAAACTGGA AAAGG GCGG	1046	
Q	y		843	A	843			
D	b		1047	A	1047			

RESULT 9
US-10-198-846-5869/c
; Sequence 5869, Application US/10198846
; Publication No. US2003099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MEI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5869
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/key: misc_feature
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

	Query Match	7.3%;	Score 277.4;	DB 14;	Length 625;
	Best Local Similarity	99.6%;	Pred. No. 5.9e-70;		
	Matches 278;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	2699	GTACTCGGCACTGAAGGCTGCACCTCGCATCAACCCCACTCAATGGTAGTGTCTGTGAAA	2758		
Db	340	GTACTCGGCACTGAAGGCTGCACCTCGCATCAACCCCACTCAATGGTAGTGTCTGTGAAA	281		
QY	2759	GGCTCGCAAAACCAACAGTCTAAGCAGTGC CGGACACCACTATGCCCTTCAGACACAGCATGTG	2818		
Db	280	GGCTCGCAAAACCAACAGTCTAAGCAGTGC CGGACACCACTATGCCCTTCAGACACAGCATGTG	221		
QY	2819	GAGATTGCACCCAGCGGACAGCTCTGAGTGCATGTGTCAGCAACATGAACGAGTGTGTGG	2878		

Db 220 GAGATTGCACCGGCGAGCTCTGAGTGATGTGGTGGAGCAACATGAAGCAGTGTGG 161
QY 2879 ACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 2938
Db 160 ACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 101
QY 2939 GCACCTGCCCCCTGAAAATGTTGAGGCTACTGTACCT 2977
Db 100 GCACCTGCCCCCTGAAAATGTTGAGGCTACTGTACCT 62

RESULT 10

US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomicca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15900
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9

US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomicca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15900
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9

Query Match 6.5%; Score 249.2; DB 9; Length 531;
Best Local Similarity 96.9%; Pred. No. 9.6e-62;
Matches 254; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1195 TTATCAAGGATATAATTTTACATGATGGAGGAAAAATTCATCCAACTGGGAATGTGACC 1254
Db 141 TTTTCTAGGATATAATTTTACATGATGGAGGAAAAATTCATCCAACTGGGAATGTGACC 200
QY 1255 AATGAGTTGAGAGTCTTTTTCACATTCATATGAGTCTGATGGTGTGTTGACCCCTAAGGCA 1314
Db 201 AATGAGTTGAGAGTCTTTTTCACATTCATATGAGTCTGATGGTGTGTTGACCCCTAAGGCA 260
QY 1315 AAGGAGCAGTATGATGATGGTCTGACACTCTGACACATTTTACACTGAAGAAATGCCGA 1374
Db 261 AAGGAGCAGTATGATGATGGTCTGACACTCTGACACATTTTACACTGAAGAAATGCCGA 320
QY 1375 GTGTCATGCTGCTGATCTTTTGTCTACTGCTCTCTATGGATATATAGCAATGTGCAG 1434
Db 321 GTGTCATGCTGCTGATCTTTTGTCTACTGCTCTCTATGGATATATAGCAATGTGCAG 380
QY 1435 GAATATGATTTGGATAAGACA 1456
Db 381 GAATATGATTTGGTAGGTATA 402

RESULT 11

US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Iodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 6.4%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2515 CTGCGAATATGCGAGTCACTCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATATGCGAGTCACTCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187
QY 2575 CTTGGAAGATCAATGTGTCTTCTACTGTGTGGGAAGATATGCCCATTTTACAAATAGT 2634
Db 186 CTTGGAAGATCAATGTGTCTTCTACTGTGTGGGAAGATATGCCCATTTTACAAATAGT 127
QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTATCTGTGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTATCTGTGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 7

```
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 12
US-09-902-941-966/c
; Sequence 966, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-966

Query Match 6.4%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCAACCTCAATGCTAGTGTCTGT 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCAACCTCAATGCTAGTGTCTGT 187

QY 2575 CTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCCATTACAAATAGT 2634
Db 186 CTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCCATTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGTCTGAGCCAGTGTCTGAGCCAGTGTCTGGAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCCGTCTGAGCCAGTGTCTGGAATTTTATCAGAA 67

QY 2695 CCAGTACTCGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2754
Db 66 CCAGTACTCGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 13
US-09-849-626-966/c
; Sequence 966, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
```

```
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-966

Query Match 6.4%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCAACCTCAATGCTAGTGTCTGT 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCAACCTCAATGCTAGTGTCTGT 187

QY 2575 CTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCCATTACAAATAGT 2634
Db 186 CTTCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCCATTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGTCTGAGCCAGTGTCTGAGCCAGTGTCTGGAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATGCCGTCTGAGCCAGTGTCTGGAATTTTATCAGAA 67

QY 2695 CCAGTACTCGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2754
Db 66 CCAGTACTCGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7

QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 14
US-10-017-754-966/c
; Sequence 966, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-966

Query Match 6.4%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCAACCTCAATGCTAGTGTCTGT 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCAACCTCAATGCTAGTGTCTGT 187
```

```
QY 2575 CTTGGAAGATCAATGTGCTTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 2634
Db |||||
186 CTTGGAAGATCAATGTGCTTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 127
QY 2635 TTACTACAGTGGATGCGCTGCTGAGCCCGAGTGCATGCTGGATTCCTGGGAATTTATCAGAA 2694
Db |||||
126 TTACTACAGTGGATGCGCTGCTGAGCCCGAGTGCATGCTGGATTCCTGGGAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db |||||
66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 2755 GAAAGG 2760
Db |||||
6 GAAAGG 1
```

RESULT 15

```
US-10-113-872-966/c
; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-966
```

```
Query Match 6.4%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCAATCTCAGAGCATGTCACGCTTAAACCCATGGTGGC 2574
Db |||||
246 CTGCGAATAATGCAGTCAATCTCAGAGCATGTCACGCTTAAACCCATGGTGGC 187
QY 2575 CTTGGAAGATCAATGTGCTTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 2634
Db |||||
186 CTTGGAAGATCAATGTGCTTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 127
QY 2635 TTACTACAGTGGATGCGCTGCTGAGCCCGAGTGCATGCTGGATTCCTGGGAATTTATCAGAA 2694
Db |||||
126 TTACTACAGTGGATGCGCTGCTGAGCCCGAGTGCATGCTGGATTCCTGGGAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db |||||
66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 2755 GAAAGG 2760
Db |||||
6 GAAAGG 1
```

Search completed: March 5, 2004, 13:18:59
Job time : 857.236 secs

2523 AATGACATCTCAGAGCATGTCACAGCTCACTTAACCCCATGGTGGGCTCGGAA 2582
Db
2520 AATCAGTATCTCAGAGCATGTCACAGCTCACTTAACCCCATGGTGGGCTCGGAA 2579
Qy
2583 GATCAATGTGCTTACTGCTGGGAGATATGTCCTCCATTTACAAATAGTTACTACA 2642
Db
2580 GATCAATGTGCTTACTGCTGGGAGATATGTCCTCCATTTACAAATAGTTACTACA 2639
Qy
2643 GTGATGCGCTCAGCCAGTATGCTGGATTTCTGGAATTTATCAGAACCCAGTAC 2702
Db
2640 GTGATGCGCTCAGCCAGTATGCTGGATTTCTGGAATTTATCAGAACCCAGTAC 2699
Qy
2703 TCGGGACTGAAGCTGCAACTGCTCACTCAACCCACTCAATGGTAGTGTCTGGAAGGCC 2762
Db
2700 TCGGGACTGAAGCTGCAACTGCTCACTCAACCCACTCAATGGTAGTGTCTGGAAGGCC 2759
Qy
2763 TCGAAACCAAGTCTAAGAGTGGGAGACCAATGTCCTTGAGGACAGCATGTGAGA 2822
Db
2760 TCGAAACCAAGTCTAAGAGTGGGAGACCAATGTCCTTGAGGACAGCATGTGAGA 2819
Qy
2823 TTGCACAGCGCAGCTCTGAGTGCATGCTGTCAGCAACATGAAGCAGTGTGTGACTC 2882
Db
2820 TTGCACAGCGCAGCTCTGAGTGCATGCTGTCAGCAACATGAAGCAGTGTGTGACTC 2879
Qy
2883 CAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTGAATGGTATACGATGAGCAC 2942
Db
2880 CAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTGAATGGTATACGATGAGCAC 2939
Qy
2943 CTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTATTGTCCTGGAGCAACAGG 3002
Db
2940 CTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTATTGTCCTGGAGCAACAGG 2999
Qy
3003 CTGTGCTGCTGATCTGATCCAGCAATCTGGCAAGGGAATGCAATGAGGGTCTCTA 3062
Db
3000 CTGTGCTGCTGATCTGATCCAGCAATCTGGCAAGGGAATGCAATGAGGGTCTCTA 3059
Qy
3063 TAAAGGACAGTGAAGTCTTCCAGCCCTTACAGCAATTTCTATCCACAGCCCT 3122
Db
3060 TAAAGGACAGTGAAGTCTTCCAGCCCTTACAGCAATTTCTATCCACAGCCCT 3119
Qy
3123 GCTCAATTCACAGTGTCTAGAGGACAGAGATACAACTGGTCTTCACTTCACTGTCC 3182
Db
3120 GCTCAATTCACAGTGTCTAGAGGACAGAGATACAACTGGTCTTCACTTCACTGTCC 3179
Qy
3183 AGCTTGCCAAATCAACGGCCACAGTAATGCAATCAATCAGACATCTGTGAGAGTGA 3242
Db
3180 AGCTTGCCAAATCAACGGCCACAGTAATGCAATCAATCAGACATCTGTGAGAGTGA 3239
Qy
3243 GAACCTGACCAAGGACAGTCTGAGACCTGCAATCTGCTTCTACGGTATCCAC 3302
Db
3240 GAACCTGACCAAGGACAGTCTGAGACCTGCAATCTGCTTCTACGGTATCCAC 3299
Qy
3303 CAATGAGGGAATGTCAAGCTATGCAAGTGCAATGGGACCGCTCTCTGTGCAACACAA 3362
Db
3300 CAATGAGGGAATGTCAAGCTATGCAAGTGCAATGGGACCGCTCTCTGTGCAACACAA 3359
Qy
3363 CAGGGCAAGTGTCTGCAACCAAGGCGTCAAGGGGAGAGTGCAGCTATGTA 3422
Db
3360 CAGGGCAAGTGTCTGCAACCAAGGCGTCAAGGGGAGAGTGCAGCTATGTA 3419
Qy
3423 GGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTTAT 3482
Db
3420 GGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTTAT 3479
Qy
3483 TGAATATCAGTTCACCTTTAGTCTATCCAGGAAGATGCTATTTACAGCTATCAA 3542
Db
3480 TGAATATCAGTTCACCTTTAGTCTATCCAGGAAGATGCTATTTACAGCTATCAA 3539
Qy
3543 TTTTGTGGTACTCTCTGCAACCAAGGATTTGGACATGTTTCAATGCTCTCAA 3602
Db
3540 TTTTGTGGTACTCTCTGCAACCAAGGATTTGGACATGTTTCAATGCTCTCAA 3599
Qy
3603 GAATTTCAACCTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGAGA 3662

3600 GAATTTCAACCTCAACATCACTCTGGGCTGCCAGTTTCTCAGTGAACCCAGGCTGGAGA 3659
Qy
3663 AGAGATGCTGTTGTTTCAAAAACCAAAATTAAGAGTACAAAGATAGTTTCTCTAATGA 3722
Db
3660 AGAGATGCTGTTGTTTCAAAAACCAAAATTAAGAGTACAAAGATAGTTTCTCTAATGA 3719
Qy
3723 GAAGTTTGATTTTTCGCAACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTTCA 3782
Db
3720 GAAGTTTGATTTTTCGCAACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTTCA 3779
Qy
3783 CTGGCCCATCAAAATTCAGTGCACCAATGAACAATGA 3819
Db
3780 CTGGCCCATCAAAATTCAGTGCACCAATGAACAATGA 3816

RESULT 2
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14

Query Match 96.8%; Score 3696.2; DB 3; Length 8589;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 3744; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
Qy 3 GGTGGCGCAGCGCGCAACTGAGGCAAGCTGAGGAGGAGGACCGCGCGCAGCGCAGC 62
Db 9 GGTGGCGCAGCGCGCAACTGAGGCAAGCTGAGGAGGAGGACCGCGCGCAGCGCAGC 68
Qy 63 GCTCGCGGCGAGGAGCGCGCGCGCGCACCGACCGCTGCGACCTGGGACGTGGAGGCC 122
Db 69 GCTCGCGGCGAGGAGCGCGCGCGCGCACCGACCGCTGCGACCTGGGAGGCC 128
Qy 123 GGGGCTGGGGCGCGGCTGCGCTCCGCGGCTGCTCTCCACCGCTGCGGCGCACCGCT 182
Db 129 GGGACCGCGCGCGCGCTGCTCTCCGCGGCTGCTCTCCGCGGCTGCTCCCGCGCGCC 188
Qy 183 GCTCTGCTGCTGTTGTTGCTCCGCGCGCGCTGTTGCTGCTGCTGCTGCTGCTGAGGC 242
Db 189 GCTCTGCTGCTGCTCTT-----TTGCTGCTGCTGCTGCTGCTGCTGCTGAGGC 239
Qy 243 CGAGGCG 302
Db 240 CGAGGCGCGCTGCGGCG 299
Qy 303 TGACCGCGCGCTGCTCAACGCGCGCTGCTCAACCGCTGCGACCGCGCGCGCGCGCGCGCG 362
Db 300 TGACCGCGCGCTGCTCAACGCGCGCTGCTCAACCGCTGCGACCGCGCGCGCGCGCGCG 359
Qy 363 CGCGCGCTGGTGGCGAGCAATGCCAGCACTGGGGGCGCGCTTACAGTAACTGATC 422
Db 360 CGCGCGCTGGTGGCGAGCAATGCCAGCACTGGGGGCGCGCTTACAGTAACTGATC 419
Qy 423 TTCTGGGTTTGTGACAGATGGACCTGGAAATTTATATAATACAAACGAAGTGCAGTGGCT 482

411 ----- 410
Db 613 TGACGAAAGGACAGCTAGAGTCTTTTGGGAAATTTTACGTGACTAAATTTTCACC 672
Qy 411 ----- ACTAATCTGATCTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAACAAA 466
Db 673 GAGAACTAACTGGCTCTCTCGAATTTGTAAACAGATGGACCTGGAAATTAATAAAGA 732
Qy 467 CGAAGTGCACGTGGCTCAATTAAGGACAGCCAAATAGAATAATGACATCTGTTCAATC 526
Db 733 CGAAGTGCACATGGCTCAATTAAGGACAGCCAAATAGAATAATGACATCTGTTCAACC 792
Qy 527 ATTTTGTCTACAGAGTGTAGTTGGGACCAATTAATGTTTATGATGGGACTCAATTTATG 586
Db 793 ATTTTGTCTACAGAAATGTAGCTGGACCAATTAATGTTTATGATGGGACTCAATCTAG 852
Qy 587 CACCGTATGTTCTGATTTAGTTGGCTCAATTTGTTCTCTGAGAGAGATGAGAGCTG 646
Db 853 CACCTCTGATTTGCTGGCTTTAGTGGCTCAATTTGTTCTCTGAGAGAGATGAGAGCTG 912
Qy 647 TCCTCTGAGTTGTTGCCACATCAGGTTATGCTTGTCTGCAATTTTATGATGGGACTCAATTTATG 706
Db 913 CTCTCTGAGTCTGCTGCTCACTTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
Qy 707 ATAAATTTGACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 766
Db 973 ATAAATCTGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1032
Qy 767 GCGAGGACAGTGTAGATCAGTATAGCAGGAACTGTTCAATGTAATGTAATGTAATGTAAT 826
Db 1033 GCGAGGACAGTGTAGATCAGTATAGCAGGAACTGTTCAATGTAATGTAATGTAATGTAAT 1092
Qy 827 ACTGGAAGGTTGAAGCATGTGATCTCTCACTGTACAGACAACTGCTGCTTCTCTCATC 886
Db 1093 ACTGGAAGGTTGAAGCATGTGATCTCTCACTGTACAGACAACTGCTGCTTCTCTCATC 1152
Qy 887 GAGGACATCTGCAATTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 946
Db 1153 GAGGACATCTGCAATTCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
Qy 947 CTGGATGTTCACTCTCTGACAGCTTAACCACTGATTTTGGACTGAGAGAAATTTCTA 1006
Db 1213 CTGGATGTTCACTCTCTGACAGCTTAACCACTGATTTTGGACTGAGAGAAATTTCTA 1272
Qy 1007 ACTTAAAGCTCCCAAGACATCTCAATAAGCTGTGGTCAATGGAACAAATTAATGTTGG 1066
Db 1273 ATTTAAAGCTCCCAAGACATCTCAATAAGCTGTGGTCAATGGAACAAATTAATGTTGG 1332
Qy 1067 TTGGAGGATATATGTTCAACCACTCAGATTAATCAATGTTTCTAGCGTATGACCTTCTT 1126
Db 1333 TTGGAGGATATATGTTCAACCACTCAGATTAATCAATGTTTCTAGCGTATGACCTTCTT 1392
Qy 1127 CTAGGAGTGTCTTCCACTAAACCGTCTGTGAACAAATGTTGGTTGTTAGATATGTTCAAT 1186
Db 1393 CTAGGAGTGTCTTCCACTAAACCGTCTGTGAACAAATGTTGGTTGTTAGATATGTTCAAT 1452
Qy 1187 CTTTGGCAATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1246
Db 1453 CTTTGGCAATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1512
Qy 1247 ATGTGACCAATGAGTTGAGAGTTTTCACATTTCAATTAATGAGTCAATGTTGTTGACCC 1306
Db 1513 ACGTGACCAATGAGTTGAGAGTTTTCACATTTCAATTAATGAGTCAATGTTGTTGACCC 1572
Qy 1307 CTAAGGCAAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
Db 1573 CGAAGCTTAAGGATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
Qy 1367 ATGCGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426
Db 1633 CTGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
Qy 1427 ATGTGAGGAATATGATTTGGATGAAGAACACATGAGATATATTACACACCCAGGTGCCCC 1486

Db 1693 TTGTGAGGAATATGACTTTGAAAGAAACACATGAGTATATTACATACTCAGGTCCTC 1752
Qy 1487 TTGTGCAAGGGGGTTAOGGCCATAGCAGTGTTTACGACCACTAGGACCAAGGCGCTATACG 1546
Db 1753 TTGTGCAAGGGGGTTAOGGCCATAGCAGTGTTTATGATGACAGGACCAAGGCTCTGTACG 1812
Qy 1547 TTCTATGTTGGCTACAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTCTACC 1606
Db 1813 TTCTATGTTGGCTACAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTCTACA 1872
Qy 1607 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1666
Db 1873 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
Qy 1667 TGCAACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1726
Db 1933 TGCAACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
Qy 1727 ACATCTATGAGCCATGCGCCCAATGCTTCTTCCAGATTTTCCAGATTTTCCAGATTTTCCAG 1786
Db 1993 ACATCTTCCATGAGCCATGCGCCCAATGCTTCTTCCAGATTTTCCAGATTTTCCAGATTTT 2052
Qy 1787 CCTGTGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1846
Db 2053 CCTGTGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
Qy 1847 GCGATTCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 1906
Db 2113 GCGATTCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 2172
Qy 1907 TCCTCAGCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 1966
Db 2173 TCCTCAGCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 2232
Qy 1967 CTGTGTTAGCAGCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGAT 2026
Db 2233 CTGTGTTAGCAGCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGAT 2292
Qy 2027 TCCTCAGCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 2086
Db 2293 TCCTCAGCAGCAGTCTTACACACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGATGAT 2352
Qy 2087 AAAGAACTCTTGAACATGACAGTGTGACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGAT 2146
Db 2353 AAAGAACTCTTGAACATGACAGTGTGACAGCAGCAGTATGATGATGATGATGATGATGATGATGATGAT 2412
Qy 2147 ACACCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2206
Db 2413 ATACCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2472
Qy 2207 AAGCCAGATCTCTCAATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2266
Db 2473 AAGCCAGATCTCTCAATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2532
Qy 2267 GTAAACAGAAAGCAGCTGACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2326
Db 2533 GCAATTAAGAAACACAGCTGACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2592
Qy 2327 GCAATCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2386
Db 2593 GCAATCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2652
Qy 2387 GAAATCTGATTTTGAATTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2446
Db 2653 GAAATCTGATTTTGAATTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2712
Qy 2447 GTAGGAAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2506
Db 2713 GTAGGAAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2772
Qy 2507 TTAAGCAGCTGCAATTAATGAGTCAATCTCAGAGCATGCTCAAGCTCAACCTTAACCCCAT 2566

QY	483	CA	TTGAAAGGACAGCCAAATAGAATAATAGAGACTTCGTTTCAATTCATTTTGTCTACAGATG	542
Db	480	CA	TTGAAAGGACAGCCAAATAGAATAATAGAGACTTCGTTTCAATTCATTTTGTCTACAGATG	539
QY	543	TA	GTGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGC	602
Db	540	TA	GTGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGC	599
QY	603	AT	TATAGTGGCCCTCAATGTTTCTCGAGAGAGATGCAATGAGACTGTCCCTGAGTGTGTTC	662
Db	600	AT	TATAGTGGCCCTCAATGTTTCTCGAGAGAGATGCAATGAGACTGTCCCTGAGTGTGTTC	659
QY	663	CACATCAGGTTATGCTTGTCTGCAATTTTATGATGCTGTCTTATAATTTGACTGGATT	722	
Db	660	CACATCAGGTTATGCTTGTCTGCAATTTTATGATGCTGTCTTATAATTTGACTGGATT	719	
QY	723	TAATATTACTTTACAGTGTGATGTGTCCAAATACTGCTCAGGCCGAGGAGAGTGAA	782	
Db	720	TAATATTACTTTACAGTGTGATGTGTCCAAATACTGCTCAGGCCGAGGAGAGTGAA	779	
QY	783	GATCAGTAATAAGCAGCGAACTGTTTGAATCTGAATGTTCTGAAAACTGGAAGGTGAAGC	842	
Db	780	GATCAGTAATAAGCAGCGAATACTGTTTGAATGTGANTGTTCTGAAAACCTGGAAGGTGAAGC	839	
QY	843	ATGTGCAATTCCTCACTGTACAGCAAACTGTGTGTTTTCTTCATCGAGGGATCTGCAATTC	902	
Db	840	ATGTGCAATTCCTCACTGTACAGCAAACTGTGTGTTTTCTTCATCGAGGGATCTGCAATTC	899	
QY	903	AAGTGATGTCTCAGAGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTCAAGTTC	962	
Db	900	AAGTGATGTCTCAGAGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTCAAGTTC	959	
QY	963	TGTACCACTAACCAAGTCAATTTTGGACTCGAGAGGAATATTCATCTTAAGACTCTCCAC	1022	
Db	960	TGTACCACTAACCAAGTCAATTTTGGACTCGAGAGGAATATTCATCTTAAGACTCTCCAC	1019	
QY	1023	AGCATCTCATAAAGCTGTGGTCAATGGAACAATTAATGTGGTGTGTGGAGGATATATGTT	1082	
Db	1020	AGCATCTCATAAAGCTGTGGTCAATGGAACAATTAATGTGGTGTGTGGAGGATATATGTT	1079	
QY	1083	CAACCACTCAGATATTAACATGTTTCHAGGGTATGACCTTGCTCTCAGGGAGTGGCTTCC	1142	
Db	1080	CAACCACTCAGATATTAACATGTTTCTAGCGTATGACCTTGCTCTCAGGGAGTGGCTTCC	1139	
QY	1143	ACTAAACCGTTCTGTGAACCAATGTGGTGTGTAGATATGTTCACTTTGGCATTATACAA	1202	
Db	1140	ACTAAACCGTTCTGTGAACCAATGTGGTGTGTAGATATGTTCACTTTGGCATTATACAA	1199	
QY	1203	GGATAAATTTACATGTATGAGAGGAATAATGTATCCAACTGGGAATGTGACCAATGAGTT	1262	
Db	1200	GGATAAATTTACATGTATGAGAGGAATAATGTATCCAACTGGGAATGTGACCAATGAGTT	1259	
QY	1263	GAGAGTTTTTCACATTCATTAATGAGTCACTGGGTGTGTTGACCCCTAAGGCAAAAGGAC	1322	
Db	1260	GAGAGTTTTTCACATTCATTAATGAGTCACTGGGTGTGTTGACCCCTAAGGCAAAAGGAC	1319	
QY	1323	GTATGAGTGTGGGCACTCTGCAACATTTGTTTACCTGAAGAATGGCCGAGTGGTCAT	1382	
Db	1320	GTATGAGTGTGGGCACTCTGCAACATTTGTTTACCTGAAGAATGGCCGAGTGGTCAT	1379	
QY	1383	GCTGCTCATCTTTGGTCACTGCCCTCTCATGGATATATAAGCAATGTGCAGGAATATGA	1442	
Db	1380	GCTGCTCATCTTTGGTCACTGCCCTCTCATGGATATATAAGCAATGTGCAGGAATATGA	1439	
QY	1443	TTTGGTAAGAACACATGGAGTATATTTACACCCAGGGTGCCCTGTGTGCAAGGGGTTA	1502	
Db	1440	TTTGGTAAGAACACATGGAGTATATTTACACCCAGGGTGCCCTGTGTGCAAGGGGTTA	1499	
QY	1503	CGGCCATAGCAGTGTGTTACGACCATAGGACCGAGGCCCTTATACGTTTCATGTGGCTACAA	1562	
Db	1500	CGGCCATAGCAGTGTGTTACGACCATAGGACCGAGGCCCTTATACGTTTCATGTGGCTACAA	1559	

QY	1563	GCGCTTTCAGTGCCCAATTAAGTACCGGCTTCGAGATGATCTCTACCGGATATGATGTGGATAC	1622
Db	1560	GCGCTTTCAGTGCCCAATAAAGTACCAGGCTTCGAGATGATCTCTACCGGATATGATGTGGATAC	1619
QY	1623	CCAGATGTGAAGCACCATCTTTAAAGACACGCCGATTTTTCCGTTACTTTGCATCAGACACAGCTGTGAT	1682
Db	1620	CCAGATGTGAAGCACCATCTTTAAAGACACGCCGATTTTTCCGTTACTTTGCATCAGACACAGCTGTGAT	1679
QY	1683	AGTCAGTGGAAACCATGCTGGTGTTTGGGGGAAAACACACACAATGACACATCTATGAGCCA	1742
Db	1680	AGTCAGTGGAAACCATGCTGGTGTTTGGGGGAAAACACACACAATGACACATCTATGAGCCA	1739
QY	1743	TGGGCCCAATATGCTTCTCTTCAGATTTTCATGGCCTATGACATTGGCTGTGACCCGCTGGTC	1802
Db	1740	TGGGCCCAATATGCTTCTCTTCAGATTTTCATGGCCTATGACATTGGCTGTGACCCGCTGGTC	1799
QY	1803	AGTGCTTCCCAGACCTGCATCTCCACCATCATGCTCAACAGATTTGGCCATTGACAGCTCTT	1862
Db	1800	AGTGCTTCCCAGACCTGATCTCCACCATCATGCTCAACAGATTTGGCCATTGACAGCTCTT	1859
QY	1863	ACACAACAGCACCATGATGATGTGTGGTGGTTTCAATAGTCTCTCTCTDAGCGACATPCT	1922
Db	1860	ACACAACAGCACCATGATGATGTGTGGTGGTTTCAATAGTCTCTCTCTDAGCGACATPCT	1919
QY	1923	GGTATTCACTTCGGAACAGTGTGATGGCGCATCGGAGTGAAGCCGCTGTTTAGCAGCANG	1982
Db	1920	GGTATTCACTTCGGAACAGTGTGATGGCGCATCGGAGTGAAGCCGCTGTTTAGCAGCANG	1979
QY	1983	ACCTGGTATTCCGTTGTGTGGAAAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGC	2042
Db	1980	ACCTGGTATTCCGTTGTGTGGAAAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGC	2039
QY	2043	AACCTGATGAACAGAGAAGAAAGTTAAATCAGATGTTTTTCCAAAAAGAACTCTTGACCA	2102
Db	2040	AACCTGATGAACAGAGAAGAAAGTTAAATCAGATGTTTTTCCAAAAAGAACTCTTGACCA	2099
QY	2103	TGACAGATGTGACACGACACACAGATTGTTTACAGCTGTACAGCCAAACCAATGACTGGCCA	2162
Db	2100	TGACAGATGTGACACGACACAGATTGTTTACAGCTGTACAGCCAAACCAATGACTGGCCA	2159
QY	2163	CTGTGTCAATGACCATTTGTGTCCCGAGGAACACACAGCTGCTCAGAGGGCCAGATCTCCAT	2222
Db	2160	CTGTGTCAATGACCATTTGTGTCCCGAGGAACACACAGCTGCTCAGAGGGCCAGATCTCCAT	2219
QY	2223	TTTTAGTGTAGAAATTGCCCCAGGAATAACCCCATGTACTCTGTGTAAACAGAACCCAG	2282
Db	2220	TTTTAGTGTAGAAATTGCCCCAGGAATAACCCCATGTACTCTGTGTAAACAGAACCCAG	2279
QY	2283	CTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGGCCCGGAANTCAGAGATGCAAT	2342
Db	2280	CTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGGCCCGGAANTCAGAGATGCAAT	2339
QY	2343	TGCCCTCCCCGAAAATATCTGTGGCATTTGGCTGGCATTTGGTGG	2387
Db	2340	TGCCCTCCCCGAGTGGCTTTGAGGCTTTGAGGCTCATCTTTGGTGTGTGGG	2384

```

RESULT 5
US-09-24
; Sequen
; Patent
; GENERA
; APPLI
; APPLI
; TITLE
; FILE
; CURRE
; CURRE
; EARLI
; EARLI
; EARLI

```


Db 2053 CTTGTGACCGATGGTCAGTCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTG 2112
QY 1847 GCCATTACAGAGTCTTACACAGACGACGATGATGTGTCGGTGTTCATATAGTCTCC 1906
Db 2113 GCCATTACAGAGTCTTGTACACAGACGACGATGATGTGTCGGTGTTCATATAGTCTCC 2172
QY 1907 TCCTCAGCAGACATCTCTGGTATTCACTTCGGAACAGTGTGATCGCATCGGATCGGAGCG 1966
Db 2173 TCCTCAGTACGTCCTTGGTCTTACCTCGAGCGATGCGATGACACACCGCATGAGCTG 2232
QY 1967 CTTGTTTACAGCAGACCTCTGGTATTCACTTCGGAACAGTGTGATCGCATCGGATCGGAGCG 2026
Db 2233 CTTGTGTCGAGCAGACCTCTGGTATTCACTTCGGAACAGTGTGATCGCATCGGATCGGAGCG 2292
QY 2027 TCTCGTGGCGCTGGCACTGATGACAGAGAGAGTAAATCAGATGTTTTCCTCA 2086
Db 2293 CTTCTGGAGTGGCACTGATGACAGAGAGAGTAAATCAGATGTTTTCCTCA 2352
QY 2087 AAAGAACTCTTGACCATGACAGATGTGACAGCAGACAGATGTTTACAGCTGCACAGCA 2146
Db 2353 AAAGAACTCTTGACCATGACAGATGTGACAGCAGACAGATGTTTACAGCTGCACAGCA 2412
QY 2147 ACACCA 2153
Db 2413 ATACCA 2419

RESULT 6
US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-12

Query Match 35.7%; Score 1364.4; DB 3; Length 6370;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1492; Conservative 7; Mismatches 48; Indels 23; Gaps 8;

QY 2257 ATGTACTACTGTACAAAGACGACGTCGAGGAGCTGTGCGCTGGACCAAGTGCAG 2316
Db 1 ATGTACTACTGTACAAAGACGACGTCGAGGAGCTGTGCGCTGGACCAAGTGCAG 60
QY 2317 TGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCATGGCTGG 2376
Db 61 TGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCATGGCTGG 120
QY 2377 CATTTGGTTGAAACTCATGTTTGAATTTACTACTGCGAGGAGATTTATGCAATGCT 2436
Db 121 CATTTGGTTGAAACTCATGTTTGAATTTACTACTGCGAGGAGATTTATGCAATGCT 180
QY 2437 AAATTTGTTCTAGGAACCAATGCGCTTTTGGCTTCTCTTACCAACCCAGAGAGGTA 2496
Db 181 AAATTTGTTCTAGGAACCAATGCGCTTTTGGCTTCTCTTACCAACCCAGAGAGGTA 240
QY 2497 GAATTTGCTTAAGCAGCTGCGCAATTAATGCAATGTCATCTCAGAGCATGTCCAGCTCAC 2556
Db

Db 241 GAATTTGCTTAAAGCAGCTGCGCAATTAATGCAATGTCATCTCAGAGCATGTCCAAAGCTCAC 300
QY 2557 TTAACCCCATGGTTCGGCTTT---CGGAAGATCAATGTGCTTACTTGG-----TGCTGGG 2608
Db 301 TTAACCCCATGGTTCGGCTTCGGGAAGGTCAATGTGCTTACTTGGTGTGGGAAG 360
QY 2609 AAGATATGTCCCATTTTACAAATAG-TTTACTACAGTGGATGCCGTCTG-----AGCCAG 2663
Db 361 GATATGTCCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTTGGAGGCCAG 420
QY 2664 TGATGCT--GGATTTCTGTGCAATTTT-----ATCAGAACCCAGTACTCGGGGACTCAAGG 2716
Db 421 TGTGCTTGGATTTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAG 480
QY 2717 CTGCAACCTGAT--CAACCCACTCAATGTGTCTGTGAAAGGCTTCGAAACACAG 2774
Db 481 CTGCAACCTGATTCACCCACTYMAATGTGTCTGTGAAAGGCTTCGAAACACAG 540
QY 2775 TGCTAA--GCAGTGCAGACACCATGTGCTTGGAGCAGCATGTGAGATTCGACACAG 2833
Db 541 TGCTAAGCAGTGCAGACACCATGTGCTTGGAGCAGCATGTGAGATTCGACACAG 600
QY 2834 GCAGCTCTGAGTGCATGTGTGCAACACATGAAGCAGTGTGTGACTTCAATGCTCATG 2893
Db 601 GCAGCTCTGAGTGCATGTGTGCAACACATGAAGCAGTGTGTGACTTCAATGCTCATG 660
QY 2894 TGGCTCTTCCCTTTTGGCAGTGTATGGAATGTATAGGATGAGCAGCTGCCCCCTG 2953
Db 661 TGGCTCTTCCCTTTTGGCAGTGTATGGAATGTATAGGATGAGCAGCTGCCCCCTG 720
QY 2954 AAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACAGCAGTGTGGTGT 3013
Db 721 AAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACAGCAGTGTGGTGT 780
QY 3014 GTACTATCCAGCAATCTGCAAGGAAATGCAATGAGGTTCTTATAAGGACCAAG 3073
Db 781 GTACTATCCAGCAATCTGCAAGGAAATGCAATGAGGTTCTTATAAGGACCAAG 840
QY 3074 TGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAGACCCCTGCTCAATTTCA 3133
Db 841 TGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAGACCCCTGCTCAATTTCA 900
QY 3134 GCATGTGTAGGACACAGATCAACTGTGCTTTCATTTCACTGTCCAGTTCCTCAAT 3193
Db 901 GCATGTGTAGGACACAGATCAACTGTGCTTTCATTTCACTGTCCAGTTCCTCAAT 960
QY 3194 GCACGCGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCA 3253
Db 961 GCACGCGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCA 1020
QY 3254 CAGGCAAGCAGCTGCGAGACCTGCATATCTGGCTTCTTACGGTATCCCAATGAGGGA 3313
Db 1021 CAGGCAAGCAGCTGCGAGACCTGCATATCTGGCTTCTTACGGTATCCCAATGAGGGA 1080
QY 3314 AATGTCAGCATGCAAGTGCATGGGACGGCTCTGTGTGCAACCAACACAGGCGAAGT 3373
Db 1081 AATGTCAGCATGCAAGTGCATGGGACGGCTCTGTGTGCAACCAACACAGGCGAAGT 1140
QY 3374 GCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAAATC 3433
Db 1141 GCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAAATC 1200
QY 3434 GATACAGGAAACCTCTCAGAGAAATGTATATATATCTCTTCTTATGACTATCAGT 3493
Db 1201 GATACAGGAAACCTCTCAGAGAAATGTATATATATCTCTTCTTATGACTATCAGT 1260
QY 3494 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTTACAGAGTATCAATTTTGGGCTA 3553
Db 1261 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTTACAGAGTATCAATTTTGGGCTA 1320
QY 3554 CTCCTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACC 3613
Db 1321 CTCCTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACC 1380

Qy	3614	TCACATCATCCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAAGATGCCTG	3673
Db	1381	TCACATCATCCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAAGATGCCTG	1440
Qy	3674	TTGTTTCAAAAACCAACATTAAAGGATACAAAGATAGTTTCTCTAATGAGAAAGTTTGATT	3733
Db	1441	TTGTTTCAAAAACCAACATTAAAGGATACAAAGATAGTTTCTCTAATGAGAAAGTTTGATT	1500
Qy	3734	TTGCGACCAACCAGAAATATCACTTTCTTTGTTATGTCACTAAATTTCACTCGGCCCATCA	3793
Db	1501	TTGCGACCAACCAGAAATATCACTTTCTTTGTTATGTCACTAAATTTCACTCGGCCCATCA	1560
Qy	3794	AAATTTCAGGT	3803
Db	1561	AAATTTCAGAT	1570

RESULT 7
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10

Query Match	16.8%;	Score 640.2;	DB 3;	Length 1051;
Best Local Similarity	86.1%;	Pred. No. 7.3e-155;		
Matches 724;	Conservative 0;	Mismatches 108;	Indels 9;	Gaps 17;
QY	3	GGTGCCCCAGCGCGCGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGC	62	
Db	216	GGTGGCGCGCGCGCGCGACTGAGGCGGGCTGAGGGGAGCACGACGACACGACGACG	275	
QY	63	GCTGCGGGCAGGACCGCGGGGCCACTGGGACTGGGACGTGGACGTGACAGGCTGGGAGGCC	122	
Db	276	GCCTGGGCGCAGGAAGGCGAGGACACGACGACCCCTGCACCGCAGCAGGGGGCTGGAGGCC	335	
QY	123	GGGGCTGGGGGCGGGCTGGCGCTCCCGGGCTGCTGTCTCCACCGCTGCGGCCCACGGCT	182	
Db	336	GGGACCGGCGCGCCCGGCTGTGTCTCCGGGGTGTGTTCGGGGGCTGCCCGCCGCC	395	
QY	183	GCTGCTGCTGCTGTTGTTGCTCCCGCGCGCGTGTGCTGTGCTGTGCTGCTGTGAGGC	242	
Db	396	GCTGCTGCGCTGCTCTT-----TTCGCTGCTGTGCTGTGCTGCTGCCGCTGCCCGGGAGGC	446	
QY	243	CGAGCCGCGCGCGCGCGCGCGGCTGTCGGGCTCAGCGCAGCGGAGGCCAAGGAATG	302	
Db	447	CGAGCCGCTGCGGGTGGCGGCGCGGTGTCGGGCTCGGCCGAGCGGAGGCCAAGGAATG	506	
QY	303	TGACCGGCCCTGTGTCAAAGCGGGTCTGCTGCAACCCCTGGCACCGGCCAATGGCTCTGGCC	362	
Db	507	TGACCGGCCGTGTGTCAACGGCGCGCGCTGTCAACCTGSCACCGGCCAGTGGCTCTGGCC	566	
QY	363	CGCCGGCTGGGTGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATC	422	
Db	567	CACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGCTTC	626	

423	QY	TTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAAATACAAACGAACTGCACGTGGCT	482
627	Db	TTCTGGATTTGTAACAGATGGACCTGGGAATTTATAAATATAGACGAATGCACATGGCT	686
483	QY	CATTGAAGCACAGCCAAATAGAAATATAGACTTCGTTTCAATCATTTTGCTACAGATG	542
687	Db	CATTGAAGCAGCCMAATAGAAATATAGACTTCGTTCAACCAITTTTGCTACAGAAATG	746
543	QY	TAGTTGGGACCAATTAATATAGTTTATGATGGGACTCAATTTATGACCGCTAGTTCTGTC	602
747	Db	TAGCTGGGACCAATTAATATAGTTTATGATGGGACTCAATCTACGCCACTCTGATTTCTGTC	806
603	QY	ATTAGTGGCCTCATGTTTCTCGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTTGTTC	662
807	Db	CTTTAGTGGCCTCATGTTTCTCGAAGAGATGGCAATGAGACGGCTCTGAGTCACGT	866
663	QY	CACATCAGGTTATGCCCTTGCTGCATTTTTTTAGTGAATGCTGTATAATTTGATCGGATT	722
867	Db	CACCTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGTATATACTGACTGGAAT	926
723	QY	TAATATATCTTACAGTTTGTATATGTCCAAATACTGCTCAGGCCGAGGACAGATGTAA	782
927	Db	TAATATACCTTACAATTTTGACATGTGTCGAATATATGCTCAGGCCGAGGACAGATGTAA	986
783	QY	GATCAGTAAATAGCAGCGGAAACTCTTGAATGTGGAATGTCTGAAAACCTGGAAGCGTGAAC	842
987	Db	GACCAATACAGCAGCAGCGCTCTTGAGTGTGAATGTCTTGAAAACCTGGAAGCGGCGG	1046
843	QY	A 843	
1047	Db	A 1047	

RESULT 8

```

US-09-702-705-966/C
; Sequence 966, Application US/09702705
; Patent No. 6504010
;
; GENERAL INFORMATION:
;
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannich, Jane
; APPLICANT: Fan, Liqun
;
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; FOR THE DIAGNOSIS OF LUNG CANCER
;
; FILE REFERENCE: 210121.478C14
;
; CURRENT APPLICATION NUMBER: US/09/702,705
;
; CURRENT FILING DATE: 2000-10-30
;
; NUMBER OF SEQ ID NOS: 1833
;
; SOFTWARE: FastSeq for Windows version 3.0
;
; SEQ ID NO 966
;
; LENGTH: 246
;
; TYPE: DNA
;
; ORGANISM: Homo sapien
;
; US-09-702-705-966

```

	Query Match	6.4%;	Score 246;	DB 4;	Length 246;
	Best Local Similarity	100.0%;	Mismatches 0;	Indels 0;	Gaps 0;
	Matches 246;	Conservative	0;	Matches 246;	Indels 0;
QY	2515	CTCGGAATAATGCGAGTCATCTCAGAGCATGTCCAAAGTCACCTTAAACCCCATGGGTGGC	2574		
DB	246	CTCGGAATAATGCGAGTCATCTCAGAGCATGTCCAAAGTCACCTTAAACCCCATGGGTGGC	187		
QY	2575	CTTCGGAAGATCAATGTGTCTCTACTGTGTCTGGGAAGATATGTCCCATTTACAAATAGT	2634		
DB	186	CTTCGGAAGATCAATGTGTCTCTACTGTGTCTGGGAAGATATGTCCCATTTACAAATAGT	127		

QY 2635 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGATCTGTGGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGATCTGTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 9
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736, 457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 6.4%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGC 187
QY 2575 CTTCGGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAGT 2634
Db 186 CTTCGGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAGT 127
QY 2635 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Query Match 6.4%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGC 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGC 187
QY 2575 CTTCGGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAGT 2634
Db 186 CTTCGGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAGT 127
QY 2635 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTATCAGAA 2694
Db 126 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966

Query Match 6.4%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e-53;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 07:06:46 ; Search time 6893 Seconds
(without alignments)
6190.781 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 1429
Sequence: 1 MVAATAAATLRRRTAATA.....KSGAVNRKQPPAQPGTCI 1429

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US09787097/runat_26022004_063015_23757/app_query.fasta_1.1607
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097 @CGN 1 1 4337 runat_26022004_063015_23757 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_esthum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	805	56.3	3976	29	AY418587	AY418587 Homo sapi
2	252	17.6	1201	13	EX440935	EX440935 EX440935
3	227	15.9	681	12	BM783739	BM783739 K-EST0061
4	226	15.8	680	12	BM783258	BM783258 K-EST0061
5	226	15.8	680	12	BM783773	BM783773 K-EST0061
6	225	15.7	677	12	BM783718	BM783718 K-EST0061
7	224	15.7	673	12	BM783685	BM783685 K-EST0061
8	210	14.7	632	12	BM773003	BM773003 K-EST0057
9	210	14.7	632	12	BM783264	BM783264 K-EST0061
10	206	14.4	743	14	CB519923	CB519923 UI-M-GIO-
11	205	14.4	3976	29	AY418589	AY418589 Mus muscu
12	193	13.5	633	10	BE894518	BE894518 601433126
13	193	13.5	933	12	EG578679	EG578679 602824477
14	192	13.4	677	12	BM783821	BM783821 K-EST0061
15	185	12.9	917	10	BF672370	BF672370 602150605
16	171	12.0	683	12	BM783788	BM783788 K-EST0061
17	161	11.3	486	12	BM772968	BM772968 K-EST0057
18	155	10.8	962	13	EX431593	EX431593 BX431593
19	154	10.8	744	13	BQ183826	BQ183826 UI-H-EUO-
20	150	10.5	544	10	BF853962	BF853962 MR2-EN009
21	142	9.9	746	14	CB242644	CB242644 UI-CF-FNO
22	141	9.9	591	12	BI850289	BI850289 imageqc_1
23	136	9.5	643	14	CA306924	CA306924 UI-H-FTI-
24	136	9.5	991	10	BE880905	BE880905 601490448
25	135	9.4	3671	29	AY418588	AY418588 Pan trogl
26	134	9.4	560	10	BF853965	BF853965 MR2-EN009
27	134	9.4	782	14	CB521623	CB521623 UI-M-GHO-
28	133	9.3	598	10	BF853006	BF853006 MR2-EN009
29	129	9.0	739	14	CA502816	CA502816 UI-CF-FNO
30	122	8.5	503	14	CA895878	CA895878 B0195B01-
31	121	8.5	686	14	CA320165	CA320165 UI-M-FWO-
32	119	8.3	706	12	BI089332	BI089332 602853292
33	119	8.3	732	14	CD644584	CD644584 AGENCOURT
34	117	8.2	390	12	EG673763	EG673763 9010 cilia
35	116	8.1	493	10	BE259981	BE259981 601148539
36	115	8.0	399	10	BF880955	BF880955 QV1-ET018
37	115	8.0	681	12	BM981492	BM981492 UI-CF-EN1
38	113	7.9	674	12	BG563729	BG563729 602584512
39	111	7.8	400	9	AA350293	AA350293 EST57512
40	110	7.7	671	9	AI818764	AI818764 willie04.X
41	108	7.6	502	9	AL048842	AL048842 DRP2P434G
42	104	7.3	484	10	BF963698	BF963698 PM4-NN120
43	99	6.9	383	9	AI372809	AI372809 EST175353
44	99	6.9	567	14	CA874419	CA874419 K0932C03-
45	99	6.9	953	10	BF313142	BF313142 601897661
46	98	6.9	595	10	BF853681	BF853681 MR2-EN009
47	95	6.6	565	14	CB525840	CB525840 UI-M-FYO-
48	93	6.5	320	9	AI372810	AI372810 EST175354
49	92	6.4	416	14	CB767839	CB767839 AMGNNUC:T
50	89	6.2	281	10	BG015062	BG015062 QV2-GN020
51	89	6.2	483	10	BF853021	BF853021 MR2-EN009
52	86	6.0	396	9	AA350292	AA350292 EST57511
53	84	5.9	758	14	CB522182	CB522182 UI-M-GHO-
54	84	5.9	778	13	BUS31905	BUS31905 AGENCOURT
55	82	5.7	484	12	BI004303	BI004303 PMO-HN007
56	81	5.7	840	14	CB520374	CB520374 UI-M-GIO-
57	80	5.6	416	14	R87660	R87660 Yp8901.s1
58	80	5.6	567	10	BF854487	BF854487 MR2-EN009
59	79	5.5	427	10	AW579586	AW579586 MR0-HT020
60	76	5.3	434	12	BI020540	BI020540 UI-M-CGDP
61	76	5.3	435	10	AW120845	AW120845 UI-M-BH2.
62	76	5.3	435	12	BI202684	BI202684 UI-M-CGDP

FEATURES		Location/Qualifiers	
source		1..3976	
gene		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		<1..3976	
		/genes="ATRN"	
		/locus_tag="HCM6607"	
ORIGIN			
Alignment Scores:		Length: 3976	
Pred. No.:		Matches: 805	
Score:		Conservative: 0	
Percent Similarity:		100.00%	
Best Local Similarity:		100.00%	
Query Match:		Indels: 0	
DB:		Gaps: 0	
US-09-787-097-12 (1-1429) x AY418587 (1-3976)			
Qy	625	ThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThr	644
Db	1559	ACCATGATGTGTTCGGTGGTTTCAATAGTCTCTCTCAGCAGACATCTGGTATTCACC	1618
Qy	645	SerGluGlnCysAspAlaHisArgSerGluAlaAaCysLeuAlaGlyProGlyIle	664
Db	1619	TCGGAACAGTGTGATGCGCATCGAGTGAAGCGCTGTGTAGCAGCAGGACCTGGTATT	1678
Qy	665	ArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGlu	684
Db	1679	CGGTGTGTGGAAACACAGGTCGTCTCAGTGATCTCTGGGGCGCTGGCACTGATGAA	1738
Qy	685	GlnGluGlnLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCys	704
Db	1739	CAAGAGAAAGTAAATCAGATGCTTTTCCAAAAGACTCTTGACCATCAGCATGT	1798
Qy	705	AspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsn	724
Db	1799	GACCAGCACACAGATTGTTACAGTCGACAGCCACACCAATGACTGCCACTGGTCAAT	1858
Qy	725	AspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyr	744
Db	1859	GACCATTTGTGTCCTCCAGGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTAGTAT	1918
Qy	745	GluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSer	764
Db	1919	GAGATTGCCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACACAGCTCGAGGAGC	1978
Qy	765	CysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro	784
Db	1979	TGTGCCCTGACAGAACCTGCCAGTGGGCCCCGGAATCAGAGGTGCATTCGCCCTGCC	2038
Qy	785	GluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThr	804
Db	2039	GAATAATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTGAATAATTA	2098
Qy	805	AlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuAla	824
Db	2099	GCCAAAGAGAAATATGACAAATGCTAAATTTGTCTGTAGGAACCAATGCCCTTTGGT	2158
Qy	825	SerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSer	844
Db	2159	TCCTTTACACCCAGGAAGAGGTAGATTTGCTTTAACGAGCTCGAATAATGCA	2218
Qy	845	SerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnVal	864
Db	2219	TCTCAGAGCATGTCCAAAGCTCACCTTAACCCATGGTGGCTTCGGAGAGTCAATG	2278
Qy	865	SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuTrpMetPro	884
Db	2279	TCTACTGTGTGGGAGAGATGATGCCCCATTTACAAATAGTTTACTACAGTGGATGCC	2338
Qy	885	SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu	904

ALIGNMENTS

AY418587 3976 bp DNA linear GSS 17-DEC-2003
Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

AY418587.1 GI:39774547

GSS.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3976)

Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,

Adams, M.D. and Cargill, M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 3976)

Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,

Adams, M.D. and Cargill, M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

These sequences were made by sequencing genomic exons and ordering

them based on alignment.


```
Db 2339 TCTGAGCCAGTGTGCTGGATCTCTGGAAATTTTATCAGAACCCAGTACTCGGGGACTG 2398
Qy 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHis 924
Db 2399 AAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTGCAACACCAC 2458
Qy 925 SerAlaIysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSer 944
Db 2459 AGTGTAAGCAGTGC CGGACACCAATGTGCTGTGAGGACAGCATGTGGAGATTGCACAGC 2518
Qy 945 GlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyr 964
Db 2519 GGCAGCTCTGAGTGCATGTGTGTGAGCAGCAACATGAGCAGTGTGTGGACTCCAATGCCAT 2578
Qy 965 ValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProPro 984
Db 2579 GTGGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCCTGCCCCCT 2638
Qy 985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
Db 2639 GAAATTTGTTGAGGCTACTGACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGG 2698
Qy 1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
Db 2699 TGTACTGATCCAGCAATACTGGCAAGGAAATGCTAGAGGGTTCCTATAAAGGACCA 2758
Qy 1025 ValLysMetProSerGluAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
Db 2759 GTGAAGATGCTTCCAGACGCCCTACAGAAATTTCTATCCACAGCCCTCTCAATTC 2818
Qy 1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGln 1064
Db 2819 AGCATGTCTAGAGGACAGCAGATACAACTGGCTCTTCACTGCTCCAGCTTGCAA 2878
Qy 1065 CysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084
Db 2879 TGCAACGGCCACAGTAATGTCATCATCAGAGCATCTCTGAGAACTGTGAGAACCTGACC 2938
Qy 1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGly 1104
Db 2939 ACAGCAAGCATCTGGACACCTGCAATNTCTGGCTTCTACGGTGATCCCAACCAATGGAGGG 2998
Qy 1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLys 1124
Db 2999 AAATGTGAGCCATGCAAGTGCATGGCAAGGCGCTCTCTGCAACACCAACACGGGCAAG 3058
Qy 1125 CysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsn 1144
Db 3059 TGCCTTCGCCACCAAGGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAAT 3118
Qy 1145 ArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGln 1164
Db 3119 CGATACCAAGGAACCCCTCTCAGAGGACATGTTATTATCTCTTCTTATTGACTATCAG 3178
Qy 1165 PheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAla 1184
Db 3179 TTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTGTGGCT 3238
Qy 1185 ThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsn 1204
Db 3239 ACTCCCTGACGACAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCAAC 3298
Qy 1205 LeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetPro 1224
Db 3299 CTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGAGATGCT 3358
Qy 1225 ValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAsp 1244
Db 3359 GTTGTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGATTGAT 3418
Qy 1245 PheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIle 1264
```

```
Db 3419 TTTCGCAACCCCAAAATATCACTTCTTTTGTATGTCAAGTAATTTACCTGGCCCATC 3478
Qy 1265 LysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVal 1284
Db 3479 AAAATTCAGATTGCCCTTCTCTCAGCACACCAATTTATGGACCTGTGTCAGATTCTTCGG 3538
Qy 1285 ThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLysIleLys 1304
Db 3539 ACTTCTTTCAGTTGTCTCTCTTGTCTCTCTGGTGGCTGTGTGGTTGGAGATCAAA 3598
Qy 1305 GlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAla 1324
Db 3599 CAAAGTTGTGGCCCTCCAGACGTGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCC 3658
Qy 1325 SerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluLeuProProAspLeu 1344
Db 3659 AGCGGTCCCTTTGCCCTCTGTAATGTCGCTTGGAAACAGATGAGAGGCTCTCTGATCT 3718
Qy 1345 IleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsn 1364
Db 3719 ATTGGGGGAGTATAAAGACTGTTCCTCAAAACCATTCGACCTGGAGCGGTGTTTGGCAAC 3778
Qy 1365 LysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPro 1384
Db 3779 AAAGCCGCTGCTCTCTCTGTGTGTGTGTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCT 3838
Qy 1385 ProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPro 1404
Db 3839 CTTGGGCGAGTCAGTCTTGTCTGGCCAGGCGCTCTGTGACATTTCTCAGCAGATGCCG 3898
Qy 1405 IleValTyrLysGlyLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGln 1424
Db 3899 ATAGTGTACAGGAGAAGTCAGAGGCGTGAGAAACCGAAGCAGAGCCCTTCGACAG 3958
Qy 1425 ProGlyThrCysIle 1429
Db 3959 CTTGGGAGCTGCATC 3973
```

RESULT 2

BX440935

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

BP 191 91006

Email: segre@genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10212.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DF012CF10QPL&cluster=10212.f. Contact :

Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0DF012CF10QPL.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DF012L19"

/tissue_type="FETAL BRAIN"

/dev stage="fetal"
 /clone lib="Homo sapiens FETAL BRAIN"
 /notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Alignment Scores:

Pred. No.: 1.77e-193 Length: 1201
 Score: 252.00 Matches: 301
 Percent Similarity: 98.69% Conservative: 0
 Best Local Similarity: 98.69% Mismatches: 2
 Query Match: 17.63% Indels: 4
 DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BX440935 (1-1201)

QY 780 CysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCys 799
 DB 58 TGCAITGCCCTGCCGAAATATCTGTGCATTTGGTCGATTTGGTTGGAACTCATGT 117
 QY 800 LeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHis 819
 DB 118 TTGAAATTTACTACTGCCAAGGAGATTTATGACAATGCTAAATTTCTGTAGGAACAC 177
 QY 820 AsnAlaLeuLeuAlaSerLeuThrThrGlnLys--LysValGluPheValLeuLysGlnL 839
 DB 178 ATGCGCTTTTGGCTCTCTTACACCCCT-GA-TGAGGTAGTAATTTGCTTTAGCAGC 235
 QY 839 euArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyL 859
 DB 236 TCGAATAATGAGTCATCTCAGAGCATGCTCAAGCTCACCTTAACCCCATGGGTGGCC 295
 QY 859 euArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerL 879
 DB 296 TTCGAAGATCAATGTCTACTGTCTGTGGAGATATGTGCCCATTTTCAAAATAGTT 355
 QY 879 euLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluP 899
 DB 356 TACTACAGTGGATCGCTCTGAGCCAGTCATGCTGGATTTCTGTGGAATTTTATCAGAAC 415
 QY 899 roSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysG 919
 DB 416 CCAGTACTCGGGAGCTGAAGGCTGAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTG 475
 QY 919 luArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaC 939
 DB 476 AAAGGCTGCACCAACACACAGTGTAGACAGTGCAGACACCATGTGCTTGAGGACGAT 535
 QY 939 ysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysV 959
 DB 536 GTGGAGATTGCAACACGCGAGCTCTGAGTGCATGTGTGTGTCAGCAACATGAAGCAGTGTG 595
 QY 959 alAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrM 979
 DB 596 TGGATCTCCATGCCATATGTGGCTCTCTCCCTTTTGGCAGTGTATGGAATGGTATACGA 655
 QY 979 etSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluG 999
 DB 656 TGAGCACCTGCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATTCCTGGAGC 715
 QY 999 InProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluG 1019
 DB 716 AACCAAGCTGTGGCTGTGTGTACTGTATCCAGCAATCTGGCAAGGGAATTCATAGAGG 775
 QY 1019 lysTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProG 1039
 DB 776 GTTCTATAAAGGACCAAGTATGATGCTTCCGAAGCCCTTACAGGAAATTTCTATCCAC 835
 QY 1039 InProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheLeH 1059

DB 836 AGCCCTGCTCAATTCAGCATGTCTAGAGACAGCAGATACAACTGGTCTTTTCATTC 895
 QY 1059 isCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluL 1079
 DB 896 ACTGTCAGCTTGCATATGCAAGCGGCACAGTAATGTCATCAATCAGAGCATCTGTGAGA 955
 QY 1079 yscysGluAsn 1082
 DB 956 AGTGTGAGAAC 966

RESULT 3

BM783739 681 bp mRNA linear EST 05-MAR-2002

LOCUS K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5', mRNA sequence.

ACCESSION BM783739

VERSION BM783739.1 GI:19131971

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 681)

AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongseung@mail.kribb.re.kr

Plate: 34 row: H column: 05

High quality sequence stop: 681.

FEATURES

source

1..681
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-34-H05"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /cell_line="SNU-484"
 /lab_host="Top10F"
 /clone_lib="S5SNU484"
 /notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A) + RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method."

ORIGIN

Alignment Scores:

Pred. No.: 1.88e-173 Length: 681
 Score: 227.00 Matches: 227
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.89% Indels: 0
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783739 (1-681)

Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerIlysLeuThrLeuThrProTrrpVal 857
 Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC 60
 Qy 858 GlyLeuArgIlysIleAsnValSerTrrpCysThrTrrpGluAspMetSerProPheThrAsn 877
 Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGGTGGGGAAGATATGTCCCATTTTACAAAT 120
 Qy 878 SerLeuLeuGlnTrrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
 Db 121 AGTTTACTACAGTGGATCGCTGAGCCAGTCATGCTGATTCGTGGAAATTTATCA 180
 Qy 898 GluProSerThrArgGlyLeuLysAlaThrCysIleAsnProLeuAsnGlySerVal 917
 Db 181 GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACCTCAATGGTAGTGTG 240
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
 Db 241 TGTGAAAGCCTCGAACCACAGTCCTAAGCAGTCCGCGACACCATGTGCTTCAGGACA 300
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrrpCysSerAsnMetLysGln 957
 Db 301 GCATGTGGAGATTGCACCGAGCGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrrpTyr 977
 Db 361 TGTGTGGACTCCAAATGCTATGTGGCTTCCTCCCTTTGGCCAGTGTATGGATGGTAT 420
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTrrpCysThrCysSerHisCysLeu 997
 Db 421 ACGATGAGCACCTGCCCCCTGAAATTTGTTCAAGCTACTGCTAGCTGAGTCAATGCTGTG 480
 Qy 998 GluGlnProGlyCysGlyTrrpCysThrAspProSerAsnThrGlyLysGlyCysIle 1017
 Db 481 GAGCAACAGGCTGTGGTGTGTACTGATCCCAATATCTGGCAAGGGAATGCATA 540
 Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
 Db 541 GAGGGTTCTATAAGGACCACTGAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 600
 Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTrrpAsnTrrpSerPhe 1057
 Db 601 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660
 Qy 1058 IleHisCysProAlaCysGln 1064
 Db 661 ATTCACTGTCCAGCTTGCCAA 681

RESULT 4
 BM783258
 LOCUS K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5', linear EST 05-MAR-2002
 DEFINITION mRNA sequence.
 ACCESSION BM783258
 VERSION BM783258.1 GI:19131490
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 680)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.Y., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr
 Plate: 31, row: H column: 08
 High quality sequence stop: 680.
 Location/Qualifiers
 1..680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-31-H08"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /cell_line="SNU-484"
 /lab_host="Top10F"
 /clone_lbb="S5SNU484"
 /notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 50nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 1,22e-172 Length: 680
 Score: 226.00 Matches: 226
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.82% Indels: 0
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783258 (1-680)

Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerIlysLeuThrLeuThrProTrrpVal 857
 Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC 60
 Qy 858 GlyLeuArgIlysIleAsnValSerTrrpCysThrTrrpGluAspMetSerProPheThrAsn 877
 Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGGTGGGGAAGATATGTCCCATTTTACAAAT 120
 Qy 878 SerLeuLeuGlnTrrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
 Db 121 AGTTTACTACAGTGGATCGCTGAGCCAGTCATGCTGATTCGTGGAAATTTATCA 180
 Qy 898 GluProSerThrArgGlyLeuLysAlaThrCysIleAsnProLeuAsnGlySerVal 917
 Db 181 GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACCTCAATGGTAGTGTG 240
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
 Db 241 TGTGAAAGCCTCGAACCACAGTCCTAAGCAGTCCGCGACACCATGTGCTTCAGGACA 300
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrrpCysSerAsnMetLysGln 957
 Db 301 GCATGTGGAGATTGCACCGAGCGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrrpTyr 977
 Db 361 TGTGTGGACTCCAAATGCTATGTGGCTTCCTCCCTTTGGCCAGTGTATGGATGGTAT 420
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTrrpCysThrCysSerHisCysLeu 997
 Db 421 ACGATGAGCACCTGCCCCCTGAAATTTGTTCAAGCTACTGCTAGCTGAGTCAATGCTGTG 480
 Qy 998 GluGlnProGlyCysGlyTrrpCysThrAspProSerAsnThrGlyLysGlyCysIle 1017

Db 481 GAGCACACAGGCTGTGGCTGCTACTGATCCAGCATACTGCGAAGGAAATGCATA 540
 QY 1018 GluGlySerTyrIysGlyProValIysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
 Db 541 GAGGGTTCCTATATAAGACACAGTGAAGATGCGCTTCGACGCCCTACAGGAAATTCAT 600
 QY 1038 ProGlnProLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
 Db 601 CCACAGCCCTGCTCAATTCAGCATGTCTAGAGCAGCAGCATACACTGGTCTTTC 660
 QY 1058 IleHisCysProAlaCys 1063
 Db 661 ATTCAGTCCAGCTTGC 678

RESULT 5
 LOCUS BM783773 680 bp mRNA linear EST 05-MAR-2002
 DEFINITION K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
 mRNA sequence.
 ACCESSION BM783773
 VERSION BM783773.1 GI:19132005
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 680)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS

Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: D column: 02
 High quality sequence stop: 680.

FEATURES

source
 1..680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484-36-D02"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /lab_host="Top10F"
 /clone_lib="S5SNU484"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was decapped with tobacco
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
 including EcoRI site by treatment of T4 RNA ligase. The
 first strand cDNA was synthesized from oligo dt-selected
 mRNA by priming with dt-tailed vector. The dt-tailed
 vector was adjusted to have about 60nt. The cDNA vector
 was circularized with E. coli DNA ligase after digestion
 of EcoRI which site is also included in vector. An RNA
 strand converted to a DNA strand by Okayama-Berg method.
 The obtained cDNA vectors were used for transfection of
 competent cells E. coli Top10F by electroporation
 method."

ORIGIN

Alignment Scores:
 Pred. No.: 1-22e-172 Length: 680
 Score: 226.00 Matches: 226
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.82% Indels: 0
 DB: 12 Caps: 0
 US-09-787-097-12 (1-1429) x BM783773 (1-680)
 QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerIysLeuThrLeuThrProTyrVal 857
 Db 1 CAGCTCGATATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTACCCCATGGGTC 60
 QY 858 GlyLeuArgIysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsn 877
 Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGCTGGGAAGATATGTCCCAATTTACAAAT 120
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
 Db 121 AGTTTACTACAGTGGATGGCTGTGAGCCAGTGAATCTGGATTCGTGGAAATTTATCA 180
 QY 898 GluProSerThrArgGlyLeuIysAlaIaThrCysIleAsnProLeuAsnGlySerVal 917
 Db 181 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTG 240
 QY 918 CysGluArgProIleAsnHisSerAlaIysGlnCysArgThrProCysAlaLeuArgThr 937
 Db 241 TGTGAAGGCTGCAACACACAGTGTAGCAGTGGCCGACCATGTGCTTTGAGGACA 300
 QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetIysGln 957
 Db 301 GCATGTGGATTTGCCAGGCGCAGCTCTGAGTGCATGTGGTGACAGCAACATGAAGCAG 360
 QY 958 CysValAspSerAsnAlaTyrValIaSerPheProPheGlyGlnCysMetGluTrpTyr 977
 Db 361 TGTGTGACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
 QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
 Db 421 ACATGAGCAGCTGCCCCCTGNAATTTGTTAGGCTACTGTACCTGTAGTATTGCTTTG 480
 QY 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyIysGlyCysIle 1017
 Db 481 GAGCAACACAGGCTGTGGCTGGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA 540
 QY 1018 GluGlySerTyrIysGlyProValIysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
 Db 541 GAGGGTTCCTATATAAGACACAGTGAAGATGCGCTTCGACGCCCTACAGGAAATTCAT 600
 QY 1038 ProGlnProLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
 Db 601 CCACAGCCCTGCTCAATTCAGCATGTCTAGAGCAGCAGCATACACTGGTCTTTC 660
 QY 1058 IleHisCysProAlaCys 1063
 Db 661 ATTCAGTCCAGCTTGC 678

RESULT 6

LOCUS BM783773

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM783718 677 bp mRNA linear EST 05-MAR-2002
 K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',
 mRNA sequence.
 BM783718 1 GI:19131950
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 677)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: F column: 01
High quality sequence stop: 677.
Location/Qualifiers

FEATURES

1. 677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-F01"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

ORIGIN

Alignment Scores:
Pred. No.: 7, 83e-172 Length: 677
Score: 225.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 15.75% Indels: 0
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783718 (1-677)

QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
DB 1 CAGCTGCCAATAATGACGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGC 60
QY 858 GlyLeuArgGlyIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
DB 61 GGCCCTCGGAAGATCAATGTGTCTTACTGTCTGGGAGATATGTCCCCATTTACAAAT 120
QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
DB 121 AGTTTACTACAGTGGATGCGGTCTGAGCCAGTCAGTGTGGATTCTGTGGAATTTATCA 180
QY 898 GluProSerThrArgGlyLeuIysAlaIaThrCysIleAsnProLeuAsnGlySerVal 917
DB 181 GAACCCAGTACTCGGGGACTGAAGCTGCACCTCATCAACCCCACTCAATGGTAGTGTGC 240
QY 918 CysGluArgProAlaAsnHisSerAlaIysGlnCysArgThrProCysAlaLeuArgThr 937
DB 241 TGTGAAGGCTGCAACACACAGTGTGAAGCAGTCCCGACACCATGTGCTTGGAGACA 300
QY 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGln 957
DB 301 GCATGTGGAGATTGCACCGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG 360
QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyr 977
DB 361 TGTGTGGACTCCAAATGCTTGTGGCTTCTTCCCTTTTGGCCAGTGTATGGATGGTAT 420

QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
DB 421 ACAGATGACACCTGCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCATCTCTTG 480
QY 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
DB 481 GAGCAACACAGGCTGTGCTGTACTGATCCAGCAATACTGCGAAGGGAATCAATCA 540
QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
DB 541 GAGGGTTCCTATTAAGACCAAGTGAAGATGCTTCCGAAGCCCTACAGGAATTTCTAT 600
QY 1038 ProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
DB 601 CCACAGCCCTGCTCAATCCAGCATGTGTCTAGAGCAGCAGCAGATACAACTGGTCTTTC 660
QY 1058 IleHisCysProAla 1062
DB 661 ATTCATGTCACGCT 675

RESULT 7

BM783685

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

1. 673

Location/Qualifiers

Source

Alignment Scores:

Pred. No.: 5,03e-171 Length: 673
 Score: 224.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.68% Indels: 0
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783685 (1-673)

QY 838 GlnLeuArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTyrVal 857
 DB 1 CAGCTGCGAATATGAGTCATCTAGACATGTCACAGTCACCTTAACCCATGGGTC 60
 QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
 DB 61 GGCCCTTCGGAAGATCAATGTCCTACTGCTGCTGGAAGATATGTCACCAATTTACAAT 120
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
 DB 121 AGTTTACTACAGTGATGCGGTCGAGCCAGTGATGCTGGATTCTGGAAATTTATCA 180
 QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
 DB 181 GAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTC 240
 QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
 DB 241 TGTGAAGGCCCTGCAACACACAGTCTAAGCAGTCCGACACCATGTGCCCTTGAGGACA 300
 QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
 DB 301 GCATGTGAGATTGCACAGCGGAGCTGAGTCATGCTGGTGACGAAATGATGGAT 360
 QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977
 DB 361 TGTGTGGACTCCAAATGCTATGCGCTCTCCCTCTTTGGCCAGTGTATGGAATGGTAT 420
 QY 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
 DB 421 AGATGAGACCTGCGCCCTCCCTGAAATGTTTGGCTACTGTACTGTAGCTAGCTGCTG 480
 QY 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyIleCysCysIle 1017
 DB 481 GAGCAACAGGCTGTGGCTGGTGTACTGATCCAGCAATCTGCAAGGAAATGCATA 540
 QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
 DB 541 GAGGGTTCCTATAAAGGACCATGAGATGCTTCCGAGGCCCTTACAGGAAATTTCTAT 600
 QY 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
 DB 601 CCACAGCCCTCTCAATTCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTTC 660
 QY 1058 IleHisCysPro 1061
 DB 661 ATTCACTGTCCA 672

RESULT 8

BM773003 632 bp mRNA linear EST 04-WAR-2002
 LOCUS K-EST0057379 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-3-E09
 DEFINITION 5', mRNA sequence.
 ACCESSION BM773003
 VERSION BM773003.1 GI:19102618
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 632)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

TITLE

JOURNAL

COMMENT

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoem-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 3 row: E column: 09
 High quality sequence stop: 632.

FEATURES

source

1..632
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S5SNU484s1-3-E09"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Epithelial"
 /cell_line="SNU-484"
 /lab_host="Top10F"
 /clone_lib="S5SNU484s1"
 /notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was decapped with tobacco
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
 including EcoRI site by treatment of T4 RNA ligase. The
 first strand cDNA was synthesized from oligo dt-selected
 mRNA by priming with dt-tailed vector. The dt-tailed
 vector was adjusted to have about 60nt. The cDNA vector
 was circularized with E. coli DNA ligase after digestion
 of EcoRI which site is also included in vector. An RNA
 strand converted to a DNA strand by Okayama-Berg method.
 The obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F, by electroporation method.
 After analyzing and sequencing about 2,000 - 3,000
 colonies in original cDNA library, the abundant cDNAs were
 selected and amplified by PCR reaction using vector region
 primer including T7 promoter as 5' primer and NidT14 as
 3' primer. The PCR products were used as template for
 synthesis of biotinylated single stranded RNA by in vitro
 transcription reaction. The synthesized RNA probes were
 hybridized with antisense single stranded cDNAs prepared
 from original library and incubated with avidin-gel.
 After removing DNA-RNA hybrids by centrifuge, the
 subtracted cDNA libraries were constructed by
 transformation of the remaining DNA into competent cells E.
 coli Top10F, with electroporation method."

ORIGIN

Alignment Scores:
 Pred. No.: 1,06e-159 Length: 632
 Score: 210.00 Matches: 210
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.70% Indels: 0
 DB: 12 Gaps: 0
 US-09-787-097-12 (1-1429) x BM773003 (1-632)
 QY 838 GlnLeuArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTyrVal 857
 DB 1 CAGCTGCGAATATGAGTCATCTAGACATGTCACAGTCACCTTAACCCATGGGTC 60
 QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
 DB 61 GGCCCTTCGGAAGATCAATGTCCTACTGCTGCTGGAAGATATGTCACCAATTTACAAT 120
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
 DB 121 AGTTTACTACAGTGATGCGGTCGAGCCAGTGATGCTGGATTCTGGAAATTTATCA 180

Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
 Db 181 GAACCAAGTACTCGGGGACTGAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTG 240
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
 Db 241 TGTGAAGGCTTGCACCAACACACAGTCTAAGCAGTCCGACACCATGTGCTTGGAGACA 300
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTyrCysSerAsnMetLysGln 957
 Db 301 GCATGTGAGATTGCACACAGCGGAGCTCTGAGTGCATGTGTGCAGCAACATGAGCAG 360
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyr 977
 Db 361 TGTGTGGACTCCAAATGCCTATGTGGCTCTCTTCCCTTTTGGCCAGTATGGAATGGTAT 420
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
 Db 421 ACGATGAGCAGCTGCCCCCTGAAATTTGTTCCAGCTACTGTACCTGTAGTCATTCCTTG 480
 Qy 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
 Db 481 GAGCAACAGAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 540
 Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
 Db 541 GAGGGTTCTCTATAAAGGACCACTGAAGATGCCCTTCGCAAGGCCCTTACAGGAAATTTCTAT 600
 Qy 1038 ProGlnProLeuLeuAsnSerMetCys 1047
 Db 601 CCACAGCCCCCTGCTCAATTCAGCATGTGT 630

RESULT 9

BM783264

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Kim YS

Unpublished (2002)

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 32, row: A column: 05

High quality sequence stop: 632.

Location/Qualifiers

1. .632

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S55NU484-32-A05"

/sex="M"

/tissue_type="Stomach"

/cell_type="Epithelial"

/cell_line="SNU-484"

/lab_host="Top10F"

/clone_lib="S55NU484"

/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;

FEATURES

source

Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."

ORIGIN

Alignment Scores:
 Pred. No.: 1.06e-159 Length: 632
 Score: 210.00 Matches: 210
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.70% Indels: 0
 DB: 12 Gaps: 0
 US-09-787-097-12 (1-1429) x BM783264 (1-632)
 Qy 838 GlnLeuArgIleValMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyrVal 857
 Db 1 CAGCTGGGAATATGACGTATCTCAGAGCATGTCGAAGCTCACCTTAACCCCATGGGTC 60
 Qy 858 GlyLeuArgLysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsn 877
 Db 61 GGCCTTCGGAAGATCAATGTGTCTACTGTGTGCGGAGATATGCCCCATTTACAAAT 120
 Qy 878 SerLeuLeuGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyLysLeuSer 897
 Db 121 AGTTTACTACAGTGGATGCCGTCTGAGCCAGCATGCTGGATCTGTGGAAATTTATCA 180
 Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
 Db 181 GAACCAAGTACTCGGGGACTGGAAGCTGCACTGCACTCAACCCACTCAATGGTAGTGT 240
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
 Db 241 TGTGAAGGCTTGCACCAACACAGTGTAGCAGTCCGACACCATGTGCTTGGAGACA 300
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTyrCysSerAsnMetLysGln 957
 Db 301 GCATGTGAGATTGCACCGCGGAGCTCTGAGTGCATGTGGTGACGACCAACATGAGCAG 360
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyr 977
 Db 361 TGTGTGACTCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
 Db 421 ACGATGAGCAGCTGCCCTTCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCATTCCTG 480
 Qy 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
 Db 481 GAGCAACAGAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 540
 Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
 Db 541 GAGGGTTCTCTATAAAGGACCACTGAAGATGCCCTTCGCAAGGCCCTTACAGGAAATTTCTAT 600
 Qy 1038 ProGlnProLeuLeuAsnSerMetCys 1047
 Db 601 CCACAGCCCCCTGCTCAATTCAGCATGTGT 630

RESULT 10

CBS19923

LOCUS

DEFINITION

UI-M-G10-cek-g-05-0-UI_r1 NIH BMAP_G10 Mus musculus cDNA clone

IMAGE:6840702 5', mRNA sequence.

CBS19923 743 bp mRNA linear EST 09-JUL-2003


```

ACCESSION      CB519923
VERSION        CB519923.1
KEYWORDS       GI:29353278
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus

REFERENCE
1 (bases 1 to 743)
NIH-MGC http://mgc.nci.nih.gov/
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: csapps@mail.nih.gov
                Tissue Procurement: Dr. Jim Lin, University of Iowa
                cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Distribution information can be found at
                http://genome.uiowa.edu/distribution/mousefl.html
                This clone was contributed by the Brain Molecular Anatomy Project
                (BMAP)
Seq primer: pYX-5.
                Location/Qualifiers
                1..743
                /organism="Mus musculus"
                /mol_type="rRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /clone="IMAGE:6840702"
                /tissue_type="whole brain"
                /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                /lab_host="DH10B (T1 phage resistant)"
                /clone_lib="NIH_BMAP_G10"
                /note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;
                Site_2: Not I; The library was constructed according
                to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                1996. Denatured RNA was size fractionated on a 1% agarose
                gel. First strand cDNA synthesis was primed with oligo-dT
                primer containing a Not I site. Double strand cDNA was
                size selected according to mRNA size fraction, ligated
                with EcoR I adaptor, digested with NotI and then cloned
                directionally into pYX-Asc vector. The library tag
                sequence located between the Not I site and the polyA tail
                is AGCGACACAG. This library was created for the University
                of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
                Developing Mouse Nervous System', supported by National
                Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                program coordinator."

ORIGIN
Alignment Scores:
Pred. No.:      2,17e-156      Length:      743
Score:          206.00      Matches:      206
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    14.42%      Indels:      0
DB:             14          Gaps:         0

US-09-787-097-12 (1-1429) x CB519923 (1-743)

QY 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
DB 64 CTGTGTTTCAAAACCAACATCAAGGATCAAGATAGTCTCTCTAATGAGAAATTT 123

QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTriPro 1263
DB 124 GATTTTGGCAACATCCAAACATCATCTTCTTTGTTTATGTCAGTAATTTCACTTGGCCC 183

QY 1264 IleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePhe 1283
DB 184 ATCAAAATTCAGATTGCCTTCTCCAGCACACGAACTTCATCGACCTCGTACAGTTCTTC 243

QY 1284 ValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTriLysIle 1303

```

```

DB 244 GTCACTTCTCTAGTTGTTTTCCTCTCGCTGCTTCTCGTGGCTGTCAGTGGTCTGGAAGATC 303
QY 1304 LysGlnSerCysTrpAlaSerArgArgGluGluGlnLeuLeuArgGlnMetGlnGlnMet 1323
DB 304 AGCAGAGACTGTTGGGCATCCAGCGGAGAGACAACTTCTTCGGAGATGCACAGATG 363
QY 1324 AlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProAsp 1343
DB 364 GCCAGCGCCCTTTGCTTCTGTAAACGTTGCTTGGAAACAGATGAAAGAACCTCTGAT 423
QY 1344 LeuIleGlyCysIleLysThrValProLysProIleAlaLeuGluProCysPheGly 1363
DB 424 CTATTGCGGGAAGTATAAGACCGTTCCTAAACCCATGCTCGGAGCCCTGCTTGT 483
QY 1364 AsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyIlePro 1383
DB 484 AACAAAGCCGACGCTCTCTCTGTATTCTGTAGGCTCCCTCGAGGACTGGAGGAATCCCT 543
QY 1384 ProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMet 1403
DB 544 CCTCCTGGTCAGTCAGGCTCTCGCTGTGGCCAGTGGCCCTGTGACATTTCTCAGCAGATG 603
QY 1404 ProIleValTyrLysGluLysSerGlyValAlaValArgAsnArgLysGlnProProAla 1423
DB 604 CCATAGTGTACAGGAGAGTCAAGGAGTGTAGAAACCGGAGCAGCAGCGCCCTGCA 663
QY 1424 GlnProGlyThrCysIle 1429
DB 664 CAGCCTGGAACCTGCATT 681

RESULT 11
AY418589 3976 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
GENOMIC SURVEY SEQUENCE.
ACCESSION AY418589
VERSION AY418589.1
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3976)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Parriera,S., Wang,G., Zheng,X.H., White,T.J., Sminsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios.
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3976)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Parriera,S., Wang,G., Zheng,X.H., White,T.J., Sminsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source 1..3976
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3976
/genes="ATRN"
/locus_tag="HCM6607"

gene
ORIGIN
Alignment Scores:

```

Pred. No.:	1.11e-155	Length:	3976
Score:	206.00	Matches:	576
Percent Similarity:	96.00%	Conservative:	0
Best Local Similarity:	96.00%	Mismatches:	12
Query Match:	14.42%	Indels:	24
DB:	29	Gaps:	0

US-09-787-097-12 (1-1429) x AY418589 (1-3976)

Qy	842	MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys	861
Db	2210	ATGCAATCATCTCAAGATATGCTCAAGCTCTGACTCCATGGGTGGTCTTCGGAAG	2269
Qy	862	IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln	881
Db	2270	ATCAATGCTTACTGCTGTGGAGGATATGCTCCATTCAAAATAGTTTCTGCAG	2329
Qy	882	TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr	901
Db	2330	TGGATGCCATCTGAGCCAGTGATGCTGGCTTCTGTGGGATCTTGTGAGACCTAGTACT	2389
Qy	902	ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro	921
Db	2390	CGGGATTAAAGGCTGCAACCTGCAACCTCTCAATGGCAGCGCTCTGGAAGGCGCT	2449
Qy	922	AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp	941
Db	2450	GCAACACACAGTCCCAAGCAGTGGCGGACACCATGTGCCCTGGGACAGCGTGGCGA-	2508
Qy	942	CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS	961
Db	2509	GTGACTAGCAG-CAGCTCGGAGTGCATGTGGTGAGTAAATGATGAAGCAGTGTGGACT	2567
Qy	961	erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerT	981
Db	2568	CCATATGCTACTGTGCT	2627
Qy	981	hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro	1000
Db	2628	G-CTGCCACCTGAAATATGCTCTGGTACTGTACTGTACTGTACTGTACTGTACTGTACT	2686
Qy	1001	GlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluLysSer	1020
Db	2687	GGCTGTGGTGTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT	2746
Qy	1021	TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln	1039
Db	2747	TATAAGGACCTGTGAAGATGCGTCACAGGCTC-TG-CAGGAAATGT-GTATCCACAG	2803
Qy	1040	ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis	1059
Db	2804	CCCTTCTGAATCCAGCATGTGTCTAGAGGACAGCATCAACTGGTCTTTTCAATCAC	2863
Qy	1060	CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys	1079
Db	2864	TGTCCAGCTTGCAGTGCACCGACACAGCAATGCATCAACACAGATATCTGTGAGAAG	2923
Qy	1080	CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs	1099
Db	2924	TGTGAGGA-CCTGACACCGGCAAGCAGCTGCGAGACCTGCATATCTGGCTTCTATGGTGA	2982
Qy	1099	pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs	1119
Db	2983	CCGACTAATGGAGGCAATGTACGCCATGCAAGTGCATGGGACGCACTCTGTGCA	3042
Qy	1119	nThrAsnThrGlyLysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe	1139
Db	3043	CACCAACACCGGCAAGTCTTCTGTACCAACCAAGAGTGTCAAGGGGACAGTGGCAGCT	3102
Qy	1139	uCysGluValGluAsnArgTyrGlnCysAsnProLeuArg-GlyThrCysTyrTyrThrL	1159
Db	3103	ATGTGAGGTAGAAATCGATACAGGGAACCTCTCAA-AGGAACATGCTACTATATCCC	3161

Qy	1159	euleuleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA	1179
Db	3162	TTCTCATTCATCATCAGTTTACCTTTAGCTGTCCAGGAAGACGACCGCTACTACACAG	3221
Qy	1179	laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA	1199
Db	3222	CCATCAACTTTGTGGCTACTCTCTGATGAACAAACAGGGATTGGACATGTTTCATCATG	3281
Qy	1199	laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl	1218
Db	3282	CTTCCCAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCGGAACCCA	3339
Qy	1218	nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer	1237
Db	3340	GAC-TGGAGAAGAGGT-GGCTGTGTGTTTCAAAACCAACATCAAGGAATACAAAGATAG	3397
Qy	1238	PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal	1257
Db	3398	TTCTCTAATGAGAAATTTGATTTTCCAAACCATCCAAACATCATCTTCTTCTTTATGTC	3457
Qy	1258	SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet	1277
Db	3458	AGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTCCAGCACAGCAACTTCATG	3517
Qy	1278	AspLeuValGlnPhePheValThrPheSerCysPheLeuSerLeuLeuValAla	1297
Db	3518	GACCTGGTACAGTCTTCGTGACTTTTTCAGTTGTTTCTCTCGCTGCTTCTGGTGGCT	3577
Qy	1298	AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeu	1317
Db	3578	GCAATGGTCTGGAAGATCAAGCAGAGCTGTGGGCATCCAGGCGGAGAGCAACTCTT	3637
Qy	1318	ArgGluMetGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr	1337
Db	3638	CGGAGATGCAACAGATGGCCAGCCCTTTGCTTCTGTAAACCTTGCCTTTGGAACA	3697
Qy	1338	AspGluGluProAspLeuIleGlySerIleLysThrValProLysProIleAla	1357
Db	3698	GATGAGAACTCTCTGATCTCATTTGGGGAGATTAAGACCTTCTTAAGCCCATGCC	3757
Qy	1358	LeuGluProCysPheGlyAsnLysAlaValLeuSerValPheValArgLeuProArg	1377
Db	3758	CTGGAGCCCTGCTTGGTAAACAAAGCCGACGCTCTCTGTATTCTGTAGGCTCCCTCGA	3817
Qy	1378	GlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal	1397
Db	3818	GGATGGAGGAATCCCTCTCTCTGTCAGTCAGGCTCTGCTGTGGCCAGTGCCTGGTG	3877
Qy	1398	AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg	1417
Db	3878	GACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGAGCTGTGAAGAACCGG	3937
Qy	1418	LysGlnGlnProProAlaGlnProGlyThrCysIle	1429
Db	3938	AAGCAGAGCGGCTGCACAGCCTGGAACCTGCATT	3973

RESULT 12
 BE894518
 LOCUS 601433126F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3918271 5',
 DEFINITION mRNA sequence.
 BE894518
 VERSION BE894518.1 GI:10356969
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 633)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/BTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9746 row: c column: 08
 High quality sequence stop: 632.
 Location/Qualifiers
 1. .633
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3918271"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 73"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 Kb. Library constructed by Life Technologies."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 6 42e-146 Length: 633
 Score: 193.00 Matches: 208
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 13.51% Indels: 1
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE894518 (1-633)

QY 1055 TrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGln 1074
 DB 6 TGGTCTTTTCATCTCACTGTGCGCAATGCCAAGCGCCACAGTAATGCAATCAATCAG 65
 QY 1075 SerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSer 1094
 DB 66 AGCATCTGTGAAGTGTGAGAACTGTGACCAGCAAGCAAGCTGTGAGCACTGTGATATCT 125
 QY 1095 GlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHis 1114
 DB 126 GGCCTTCTACGTGATCCCAATGAGGGAATGTCCAGCAATGCAAGTCAATGGGCAC 185
 QY 1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134
 DB 186 GGTCTCTGTGCAACCAACACAGCGGCAAGTCTTGTGACCAACCAAGGGGTCAAGGGG 245
 QY 1135 AspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThr 1154
 DB 246 GACGAGTGCACCTATGTGAGTAGAATCGATACCAAGGAACTCTCTCAGAGGAACA 305
 QY 1155 CysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAsp 1174
 DB 306 TGTATTATATCTCTTTATTGACTATCATGTTCACTTGTAGTCTATCCAGGAAGATGAT 365
 QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAsp 1194
 DB 366 CGCTATTACACAGCTATCAATTTTGGGTACTCTCTGACCAACCAACAGGGAATTTGGAC 425
 QY 1195 MetPheIleAsnAlaSerIleAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSer 1214
 DB 426 ATGTTTCATCAATGCGCTCCCAAGAAATTCACCTCAACATCACCTGGGTGCGGCTTCTCA 485
 QY 1215 AlaGlyThrGlnAlaGlyGluMetProValValSerLysThrAsnIleLysGluTyr 1234
 DB 486 GCTGGAAACCCAGCTGGAGAGAGATGCTGTGTTTCCAAAACCAACATTAAGAGATAC 545
 QY 1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254

Db 546 AAGATAGATTTCTCTAATGAGAAAGTTTGATTTTTCGCAA-CACCCAAATATCACTTTCTTT 604
 QY 1255 ValTyrValSerAsnPheThrTipPro 1263
 DB 605 GTTATATGTCAGTAATTTTCACTGGCCC 631

RESULT 13

BG678679

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

ORIGIN

Alignment Scores:
 Pred. No.: 9 38e-146 Length: 933
 Score: 193.00 Matches: 242
 Percent Similarity: 99.18% Conservative: 0
 Best Local Similarity: 99.18% Mismatches: 1
 Query Match: 13.51% Indels: 2
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BG678679 (1-933)

QY 1000 ProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019
 DB 3 CCAGGCTGTGGTGTGTGTACTGATCCCAATCTGCGCAAGGGAATGCATAGAGGT 62
 QY 1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
 DB 63 TCCATPAAAGGACCAAGTGAAGATGCTTCGAAGCCCTACAGGAATTTCTATCCACAG 122
 QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHis 1059
 DB 123 CCCCTGCTCAATTCACGATGTCTAGAGGACAGACAGATCACTGGTCTTTTCATTCAC 182
 QY 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
 DB 183 TGTCCAGCTTGGCAATGCAACGCGCCACAGTAATGATCAATCAGATCTGTGAGAAG 242

```

QY 1080 CysGluAenLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAsp 1099
DB 243 TGTGAGAACCTGACACAGGCAAGCACTGCAGACCTGCATATCTGGCTTCTACGGTGAT 302
QY 1100 ProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsn 1119
DB 303 CCCACCAATGGAGGAATGTCAGCCATGCAAGTCAATGGCAGCGCTCTGTGCAAC 362
QY 1120 ThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu 1139
DB 363 ACCAACCGGCGAAGTGTCTTCGCCACCAAGGCGCTCAAGGGGAGCAGTGCACGTA 422
QY 1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeu 1159
DB 423 TGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACATGTTATATACTCTT 482
QY 1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAla 1179
DB 483 CTTATTACTATCAGTTCACCTTACCTTATAGTCTATCCAGGAAGATGATCGCTATTACAGCT 542
QY 1180 IleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu-AspMetPheIleAsnAl 1199
DB 543 ATCAATTTTGTGGCTACTCTCGACGAACAAACAGGGATTC-GGACATGTTCAICATGCG 601
QY 1199 aSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAl 1219
DB 602 CTCCAAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 661
QY 1219 aGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSe 1239
DB 662 TCGAGAAGAGATGCGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 721
QY 1239 rAsnGluLys 1242
DB 722 TAATGAGAG 731

RESULT 14
BM783821
LOCUS
DEFINITION K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5',
mRNA sequence.
ACCESSION BM783821
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eeun-dong Yuseong-gu, Daejeon 305-393, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: H column: 09
High quality sequence stop: 677.
FEATURES
Location/Qualifiers
1..677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-H09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"

```

```

/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation
method."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 4,44e-145 Length: 677
Score: 192.00 Matches: 224
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 13.44% Indels: 2
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783821 (1-677)

QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
DB 1 CAGCTGCCAATAATGCAATGCTATCTCAGAGCATGCTCCNAGCTCACCTTAACCCCATGGTTC 60
QY 858 GlyLeuArgLysIleAsnValSerTyrTyrCysTrpGluAspMetSerProPheThrAsn 877
DB 61 GGCCTTCGAAGATCAATGTGCTCTACTGTGTGGAGAGATATGTCCTCCCATTTACAAAT 120
QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
DB 121 AGTTTACTAGAGTGATGCGCTCTGAGCCAGTGATGCTGGATTCTGTGGAATTTATCA 180
QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
DB 181 GAACCCAGTACTCGGGACTGGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTTC 240
QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
DB 241 TGTGAAGGGCTGCAAAACCAACACAGTGTCTAAGCAGTGGCGGACACCATGTGCTTGAGACA 300
QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
DB 301 GCATGTGAGATTGCACACGCGGAGCTCTGAGTGCATGTGGTGACAGCAACATGAACAG 360
QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyr 977
DB 361 TGTGTGACTCCAAATGCCTATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
DB 421 ACAGTAGCAGCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 480
QY 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
DB 481 GAGCAACCAAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGAAAGTGCATA 540
QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGln-AlaProThrGlyAsnPheTy 1037
DB 541 GAGGGTTCTCTATAAAGGACCAAGTGAAGATGCTCTCGNC-AGCCCCCTCAGGAAATTTCTA 599
QY 1037 rProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPh 1057
DB 600 TCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGAGCAGCAGATACAACTGGTCTTT 659
QY 1057 eIleHisCysProAla 1062

```

```

Db      660 CATTCTGTCAGCT 675
|||||
RESULT 15
BF672370      917 bp mRNA linear EST 21-DEC-2000
LOCUS      602150605F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4291680 5',
DEFINITION mRNA sequence.
ACCESSION BF672370 GI:11946265
VERSION BF672370.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 917)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM136 row: j column: 01
High quality sequence stop: 563.
FEATURES
source
1. .917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4291680"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 81"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI
(ggccattatgcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCCATATGACC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGGCCGACATG-dT(30)EN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

ORIGIN

```

Alignment Scores:
Pred. No.:      2,82e-139      Length:      917
Score:      185.00      Matches:      185
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 12.95%      Indels: 0
DB:      10      Gaps: 0
```

US-09-787-097-12 (1-1429) x BF672370 (1-917)

```

QY      945 GlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyr 964
Db      1 GGAGGCTCTGAGTGCAATGGTGGCAGCACATGAGGAGTGTGGATCCATGCCCTAT 60
QY      965 ValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 984
Db      61 GTGGCCTCTCCCTTTGGCCAGTGTATGTAATGGTATACGATGAGCACCTGCCCCCT 120
QY      985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTrp 1004
Db      121 GAAATTTGTTACGGCTACTGTACCTGATCTTGTGGAGCAACAGGCTGGCTGG 180
```

```

QY      1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
Db      181 TGTACTGATCCCGCAATACTGGCAAGGAATGCAATAGAGGGTTCCTATAAGGACCA 240
QY      1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
Db      241 GTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCC 300
QY      1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGln 1064
Db      301 AGCATGTGTCTAGAGGACAGCAGATCAACTGTGCTTTTCATTCACCTGTCAGCTTGCCAA 360
QY      1065 CysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGlyLysCysGluAsnLeuThr 1084
Db      361 TGCACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 420
QY      1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGly 1104
Db      421 ACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTTCTACGGTGTATCCCAATGGAGGG 480
QY      1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLys 1124
Db      481 AAATGTGAGCATGCAAGTCAATGGCAGCCGCTCTCTGTGCAACACACACGGGCAAG 540
QY      1125 CysPheCysThrThr 1129
Db      541 TGCTTCTGCACCACC 555

RESULT 16
BF783788      683 bp mRNA linear EST 05-MAR-2002
LOCUS      K-EST0061794 SSSNU484 Homo sapiens cDNA clone SSSNU484-36-E09 5',
DEFINITION mRNA sequence.
ACCESSION BF783788
VERSION BF783788.1 GI:19132020
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 683)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: E column: 09
High quality sequence stop: 683.
FEATURES
source
1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSSNU484-36-E09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="SSSNU484"
High quality sequence stop: 683.
```

vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F⁺ by electroporation method."

ORIGIN					
Alignment Scores:					
Pred. No.:	4.74e-128	Length:	683		
Score:	171.00	Matches:	225		
Percent Similarity:	98.25%	Conservative:	0		
Best Local Similarity:	98.25%	Mismatches:	2		
Query Match:	11.97%	Indels:	4		
DB:	12	Gaps:	0		
US-09-787-097-12 (1-1429) x BM772968 (1-683)					
QY	838	GlnLeuArglleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTIPVal	857		
DB	1	CAGTGGCAATATATGAGTATCTCAGAGCATGTCACAGCTACCTTAACCCCATGGGTC	60		
QY	858	GlyLeuArgLysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsn	877		
DB	61	GCCTTCGGAAGATCAATGTGCTCTACTGCTGGGAGATATGTCCTCCCATTTACAAAT	120		
QY	878	SerLeuLeuGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer	897		
DB	121	AGTTTACTACAGTGGATCGGCTCGAGCCAGTGTGCTGGAATTTTATCA	180		
QY	898	GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal	917		
DB	181	GAACCCAGTACTCGGGAGTCAAGCTGCACTGCACTCAACCCCATCAATGGTAGTGC	240		
QY	918	CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr	937		
DB	241	TGTGAAGGCTGCAAAACACAGTCTAAGCAGTCCGACACCATGTCCTTGAGGACA	300		
QY	938	AlaCysGlyAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLysGln	957		
DB	301	CGATGTGAGATTGACACAGCGGAGCTCTGAGTCGATGTTGTCAGCAACATGAGCAG	360		
QY	958	CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyr	977		
DB	361	TGTGTGGACTCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTATGGAATGGTAT	420		
QY	978	ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu	997		
DB	421	ACGATGAGCAGCTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCAITGCTTG	480		
QY	998	GluGlnProGlyCysGlyTyrCysThrAspProSerAsn--ThrGlyLysGlyLysCysI	1017		
DB	481	GAGCAACAGGCTGTGGCTGTGTACTGATCCAG-NC-ATAGTGGCAAGGGAATGCA	538		
QY	1017	leGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheT	1037		
DB	539	TAGAGGGTTTCTATTAAGGACAGTGAAGATGCTCTGGCAACCCCTACAGGAATTTCT	598		
QY	1037	YrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerP	1057		
DB	599	ATCCAGAGCCCTGTCTCAATTCACGATGTGTCTAGAGGACAGACATCACTGGTCTT	658		
QY	1057	heileHisCysProAlaCysGln	1064		
DB	659	TCATTCAGTGTCCAGCTTGCCAA	681		

RESULT 17
BM772968
LOCUS
DEFINITION K-ESTC057334 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-3-A12
5', mRNA sequence.
ACCESSION BM772968

BM772968.1 GI:19102583

EST. Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 486)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eeum-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 3 row: A column: 12

High quality sequence stop: 486.

Location/Qualifiers

1..486

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S5SNU484s1-3-A12"

/sex="M"

/tissue_type="Stomach"

/cell_type="Epithelial"

/cell_line="SNU-484"

/lab_host="Top10F"

/clone_lib="S5SNU484s1"

/notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F⁺ by electroporation method. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F⁺ with electroporation method."

ORIGIN

Alignment Scores:

Pred. No.: 4.35e-120 Length: 486

Score: 161.00 Matches: 161

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 11.27% Indels: 0

DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM772968 (1-486)

QY 838 GlnLeuArglleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTIPVal 857

DB 1 CAGTGGCAATATATGAGTATCTCAGAGCATGTCACAGCTACCTTAACCCCATGGGTC 60

QY 858 GlyLeuArgLysIleAsnValSerTyrThrProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 877
 Db 61 GGCCCTTCGGAAGATCAATGTCTCTACTGGTGGGAAGATATGTCCCAATTTACAAT 120
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
 Db 121 AGTTTACTACAGTGGATGGCGCTGAGCCAGTGGATCTGGATCTGGTGGATTTATCA 180
 QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
 Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAAGCTGCAATCAACCCCACTCAATGGTAGTGT 240
 QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
 Db 241 TGTGAAGGCTTGCACCAACACAGTGTCTAAGCAGTCCGACACCATGTGCTTGAAGACA 300
 QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
 Db 301 GCATGTGGAGATTGCACACGCGGAGCTCTGAGTGCATGCTGGTGAGCAACATGAAGCAG 360
 QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977
 Db 361 TGTGTGGATCTCAATGCCCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGATGGTAT 420
 QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
 Db 421 ACGATGAGCACCTGCCCTTCAATTTGTTTCCAGCTACTGTACCTGTAGTCAATGTCTTG 480
 QY 998 Glu 998
 Db 481 GAG 483

RESULT 18
 BX431593
 LOCUS
 DEFINITION
 CS0DF012YL19 5-PRIME, mRNA sequence.
 BX431593
 BX431593.1 GI:30781048
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 962)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10212.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAG0172B02_CS01546_1&cluster=10212.f.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/InvitrogenCorporation1600
 Faraday Avenue Genoscope sequence ID : CS0BAG0172B02_CS01546_1.
 Location/Qualifiers
 1. 962
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF012YL19"
 /tissue type="FETAL BRAIN"
 /dev stage="fetal"
 /clone.lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo (GT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and

FEATURES
 source

cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,19e-115 Length: 962
 Score: 155.00 Matches: 204
 Percent Similarity: 99.03% Conservatives: 0
 Best Local Similarity: 99.03% Mismatches: 1
 Query Match: 10.85% Indels: 2
 DB: 13 Gaps: 0
 US-09-787-097-12 (1-1429) x BX431593 (1-962)
 QY 1054 AsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsn 1073
 Db 53 AACTGGTCTTTCATTCCTCCAGCTTCCCAATGCAACGGCCACAGTAATGATCAAT 112
 QY 1074 GlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIle 1093
 Db 113 CAGGACATCTGTGAGAGTGTGAGAACCTGACCAAGGACAGCACTGCGAGACCTGCATA 172
 QY 1094 SerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGly 1113
 Db 173 TCTGGCTTCTACGGTGTATCCCAATGGAGGGAATGTCCAGCATGCAATGCAATGG 232
 QY 1114 HisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLys 1133
 Db 233 CACGGTCTCTGTGCACACCAACACGGGCAAGTGTCTTGACACCAAGGGCGTCAAG 292
 QY 1134 GlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGly 1153
 Db 293 GGGGACGAGTCCAGCTATGTGAGGTAGAAATTCGATACCAAGGAACCTCTCGAGGA 352
 QY 1154 ThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAsp 1173
 Db 353 ACATGTATTATTAATCTCTTCTATTGACTACTAGTTCACCTTTAGTCTATCCCAGGAAGAT 412
 QY 1174 AspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu 1193
 Db 413 GATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGACAAACACAGGATTTG 472
 QY 1194 AspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrp-AlaAlaSerPh 1213
 Db 473 GACATGTTTCATCAATGCCCTCCAGAAATTTCAACCTCAACATCACCT-GGCTGCCAGTTT 531
 QY 1213 eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGln 1233
 Db 532 CTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTGTTCATAAAACCAACATTAAAGA 591
 QY 1233 uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh 1253
 Db 592 GTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGACCAACCAACCAATATCACTTT 651
 QY 1253 ePheValTyrValSer 1258
 Db 652 CTTTGTATTATGTCAGT 667

RESULT 19
 BQ183626/c
 LOCUS
 DEFINITION
 BQ183626 744 bp mRNA linear EST 15-JUL-2003
 UI-H-EUO-azn-a-22-0-UI.s1 NCI-CGAP_Carl Homo sapiens cDNA clone
 IMAGE:5850909 3', mRNA sequence.
 BQ183626
 VERSION
 BQ183626.1 GI:20359183
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 744)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 . TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 16-237, >11P22#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers
1..744
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5850909"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Carl"
/note="Organ: Knees; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Carl is a cDNA library containing the following tissue(s): Osteoarthritic Cartilage. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.
TAG_TISSUE=osteoarthritic cartilage
TAG_LIB=UI-H-EUO
TAG_SEQ=TGATCAGCT"

ORIGIN

Alignment Scores:
Pred. No.: 3,12e-114 Length: 744
Score: 154.00 Matches: 154
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.78% Indels: 0
DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BQ183626 (1-744)

QY 1114 HisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrIysGlyVallys 1133
Db 697 CACGCGTCTGTGCAACCAACACGCGGCAAGTGTCTTGACCACCAAGGGCGTCAAG 638
QY 1134 GlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGly 1153
Db 637 GGAGACGAGTGCAGCTATGTAGGTAGAAATCGATACACAGGAACCTCTCAGAGA 578
QY 1154 ThrCysTyrThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGluAsp 1173
Db 577 ACATGTTATTACTCTCTTATTGACTATCAGTTCACCTTAGTCTATCCAGGAAGAT 518
QY 1174 AspArgTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu 1193
Db 517 GATCGTATTACACAGCTATCAATTTGTGGCTACTCTCGACCAACAAACAGGGATTTC 458
QY 1194 AspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIaSerPhe 1213
Db 457 GACATGTTTCATCATGCTTCAGAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTTC 398

QY 1214 SerAlaGlyThrGlnAlaGlyGlnGluMetProValValSerLysThrAsnIleLysGlu 1233
Db 397 TCAGCTGGAAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAGGAG 338
QY 1234 TyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhe 1253
Db 337 TACAAAGATAGTTCCTCTAATGAGAAGTTGATTTTCGCAACCAACCAATATCACTTTC 278
QY 1254 PheValTyrValSerAsnPheThrTrpProLleLysIleGln 1267
Db 277 TTTGTTTATGTGATTAATTCACCTGCCCATCAAAAATTCAG 236

RESULT 20
BF853962
LOCUS BF853962 544 bp mRNA linear EST 16-JAN-2001
DEFINITION MR2-EN0093-261200-004-g08 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853962
VERSION BF853962.1 GI:12241706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 544)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brines, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-261200-004-g08&t3=2000-12-26&t4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 544.
Location/Qualifiers
1..544
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/note="Organ: lung_normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 4.02e-111 Length: 544
Score: 150.00 Matches: 150
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.50% Indels: 0
DB: 10 Gaps: 10

FEATURES
source

US-09-787-097-12 (1-1429) x BF853962 (1-544)

```

QY 1113 GlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyVal 1132
Db 64 GGCACGCGTCTCTGTGACACCAACACGGGAGTGTCTTCACCAACCAAGGCGTC 123
QY 1133 LysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg 1152
Db 124 AAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCAGA 183
QY 1153 GlyThrCysTyrThrThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGlu 1172
Db 184 GGAACATGTTACTATCTCTCTTATGACTATCAGTTCACTTACCTTGTAGTCTATCCAGGAA 243
QY 1173 AspAspArgTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAsp 1192
Db 244 GATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTCTGACGAACAAACAGGGAT 303
QY 1193 LeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSer 1212
Db 304 TTGGATGTTTCATCAATGCTCCCAAGAAATTTCAACCTCAACATCACTGGGCTGCCAGT 363
QY 1213 PheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLys 1232
Db 364 TTCTCAGCTGGACCCAGGCTGGAGAGAGATGCTGTGTGTTTCAAAACCAACATTAAAG 423
QY 1233 GluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThr 1252
Db 424 GAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACT 483
QY 1253 PhePheValTyrValSerAsnPheThrTrp 1262
Db 484 TTTCTTTGTTATGTCAGTAATTTCACTCGG 513

```

RESULT 21

CB242644/c

LOCUS

DEFINITION

UI-CF-FNO-afs-i-02-0-UI s1 UI-CF-FNO Homo sapiens cDNA clone

UI-CF-FNO-afs-i-02-0-UI 3', mRNA sequence.

CB242644

CB242644.1 GI:28364288

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 746)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 16-237, >L1PA2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

Location/Qualifiers

1..746

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-FNO-afs-i-02-0-UI"

/tissue_type="Human Lung Epithelial cells"

/lab_hosts="DH10B (Life Technologies) (T1 phage resistant)";

/clone_lib="UI-CF-FNO"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-CF-FNO is a subtracted cDNA library derived from two

normalized human lung epithelial cell libraries (EN1 and

DUI) The library was subtracted according to according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:

bento-soares@uiowa.edu

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr to LPS 24h

TAG_LIB=UI-CF-FNO

TAG_SEQ=CTGCTCAGGT"

ORIGIN

Alignment Scores:

```

Pred. No.: 1.67e-104 Length: 746
Score: 142.00 Matches: 169
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 1
Query Match: 9.94% Indels: 1
DB: 14 Gaps: 0

```

US-09-787-097-12 (1-1429) x CB242644 (1-746)

```

QY 1098 GlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117
Db 744 GGTGATCCCAATGAGGAAATGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 685
QY 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137
Db 684 TGCACACCAACACGGGCAAGTG-TTCTGCACCAACAGGGCGTCAAGGGGACGAGTGC 626
QY 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157
Db 625 CAGCTATGTGAGGTGAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTAT 566
QY 1158 ThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyr 1177
Db 565 ACTCTTCTTATTGACTATCAGTTTCACTTTAGTCTTATCCAGGAGATGATCGCTATTATAC 506
QY 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
Db 505 ACAGCTATCAATTTTGTGGCTACTCTCTGACGACAAACAAACAGGGATTTGGACATGTCATC 446
QY 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
Db 445 AATGCCTCCAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACC 386
QY 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
Db 385 CAGGCTGGAGAGAGATGCTGTGTGTTTCAAAACCAACATTAAGAGATACAAAGTAGT 326
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
Db 325 TTCTCTTAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACTTCTTCTTTATGTC 266
QY 1258 SerAsnPheThrTrpProIleLysIleGln 1267
Db 265 AGTAATTTCACTGGCCCATCAAAATTCAG 236

```

RESULT 22

BI850289

LOCUS

DEFINITION

imageqc 11.2000/sly246bdr81.v1 NH_MGC_56 Homo sapiens cDNA clone

IMAGE:4284387 5', mRNA sequence.

BI850289

ACCESSION

BI850289

```

VERSION BI850289.1 GI:16003776
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 591)
AUTHORS Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
Prange, C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
resequencing for verification
JOURNAL Unpublished (2001)
COMMENT Other_ESTS: BF700035
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
control effort. High quality sequence is defined as having 100 or
more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs marks the beginning and end of the
sequence. For information on obtaining this clone, please contact
info@image.llnl.gov.
Plate: L10M117 row: j column: 4
Seq primer: -21m13
High quality sequence stop: 591.
Location/Qualifiers
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4284387"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH MGC 56"
/notes="Organ: Brain; Vector: pBMR-LIB (Clontech); Site 1:
SFII (ggccctcgcc); Site 2: SFII (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGGCATTAGGCC-3' and 3' adaptor
sequence: 5'-ATTTAGAGCGGCGGCGGACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

ORIGIN

```

Alignment Scores:
Pred. No.: 8,62e-104 Length: 591
Score: 141.00 Matches: 195
Percent Similarity: 99.49% Conservatives: 0
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 9.87% Indels: 1
DB: 12 Gaps: 0
```

US-09-787-097-12 (1-1429) x BI850289 (1-591)

```

QY 880 LeuGluTrpMetProSerGluProSerAspAlaGlyPheCysGlyTleuSerGluPro 899
|||||
Db 3 CTACAGTGGATGCGCTGTAGCGCCAGTGTGGATTCTGTGGAAATTTATCAGAACCC 62
|||||
QY 900 SerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGlu 919
|||||
Db 63 AGTACTCGGGGACTGAGGCTGCAACTGCATCAACCCACTCAATGTAGTGTGTGAA 122
|||||
QY 920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939
|||||
Db 123 AGGCTCTGCAACACACAGTGTACAGCAGTGTCCGACACCATGTGC-CTGAGGACAGCATGT 181
|||||
QY 940 GlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959
|||||
```

```

Db 182 GGAGATTGCACACGCGCGAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGTG 241
|||||
QY 960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMet 979
|||||
Db 242 GACTCCAAATGCCATATGTGGCCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATG 301
|||||
QY 980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999
|||||
Db 302 AGCACCTGCCCCCTGAAAATTGTTGAGGCTACTGTACCTGTAGTCTGATGCTGTGGAGCAA 361
|||||
QY 1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyGlyGlyCysIleGluGly 1019
|||||
Db 362 CCAGGCTGTGGTGTGTACTGATCCAGCAATACTGCGCAAGGGAATGATAGAGGCT 421
|||||
QY 1020 SerTyrIysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
|||||
Db 422 TCTATAAAGGACAGTGAAGATGCTTCGAGAGCCCTACAGGAATTTCTATCCACAG 481
|||||
QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
|||||
Db 482 CCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCAATTCAC 541
|||||
QY 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSer 1075
|||||
Db 542 TGTCCAGCTTGCCATGCACAGGCCACAGTAAATGCATCAATCAGAGC 589
|||||
```

RESULT 23

```

CA306924/c 643 bp mRNA linear EST 01-NOV-2002
LOCUS UI-H-FT1-bhu-e-21-0-UI.s1 NCICGAP_FTL Homo sapiens cDNA clone
DEFINITION UI-H-FT1-bhu-e-21-0-UI 3', mRNA sequence.
```

ACCESSION CA306924

VERSION CA306924.1 GI:24469978

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 643)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Gary W. Hunninghake, U of I

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA

sequence: 16-237, >L1PA2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA-Yes.

Location/Qualifiers

1..643

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-FT1-bhu-e-21-0-UI"

/tissue_type="Aveolar Macrophage"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI_CGAP_FTL"

/note="Organ: Lung; Vector: p7T73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI CGAP FTL is a normalized cDNA library constructed from

a pool of 81 RNA samples from Alveolar Macrophages

challenged with different treatments. The library was

normalized according to Bonaldo, Lemon and Soares, Genome

Research, 6:791-806, 1996. First strand cDNA synthesis was

primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCCATGCCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.

TAG_LIB=UI-H-FTI
TAG_SEQ=GGCCATGCCG

ORIGIN

Alignment Scores:
Pred. No.: 1,06e-99 Length: 643
Score: 136.00 Matches: 136
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.52% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA306924 (1-643)

QY 1132 ValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeu 1151
DB 643 GTCAAGGCGNACGAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGGAACCTCTC 584
QY 1152 ArgGlyThrCysTyrTyrThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGln 1171
DB 583 AGAGGAACATGTTATTATCTCTTCTTATGACTATCATCTTACCTTTAGTCTATCCAG 524
QY 1172 GluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArg 1191
DB 523 GAAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTCGAGCAACAAACAGG 464
QY 1192 AspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla 1211
DB 463 GATTGGAGATGTTTCATCAATCGCTCCCAAGATTTTCAACCTCAACATCACCTGGGCTGCC 404
QY 1212 SerPheSerAlaGlyThrGlnAlaGlyGluMetProValValSerLysThrAsnIle 1231
DB 403 AGTTTCTCAGCTGGAAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATT 344
QY 1232 LysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIle 1251
DB 343 AAGGAGTACAAAGATGTTCTCTAATGAGAAGTTTGATTTTGGCAACCAACCAAAATATC 284
QY 1252 ThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
DB 283 ACTTTCCTTGTGTTTGTGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 236

RESULT 24

BE880905
LOCUS
DEFINITION BE880905 991 bp mRNA linear EST 20-OCT-2000
601490448P1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3892859 5',
mRNA sequence.
ACCESSION BE880905
VERSION BE880905.1 GI:10329681
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 991)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DNP/Gazdar
CNA Library Preparation: Life Technologies, Inc.
CNA Library Arrayed by: The I.W.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.W.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM9679 row: p column: 12
High quality sequence stop: 607.

FEATURES

source
1..991
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3892859"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_69"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dr. Average insert size 1.1 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 1,61e-99 Length: 991
Score: 136.00 Matches: 176
Percent Similarity: 98.32% Conservative: 0
Best Local Similarity: 98.32% Mismatches: 1
Query Match: 9.52% Indels: 3
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE880905 (1-991)

QY 513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532
DB 39 GCCTATACGTTTCATGTGTGCTACAAAGCTTTTCAGTGCCTCAATAGTACCGCTTCGAGAT 98
QY 533 AspLeuTyrArgTyrAspValAspThrGlnMetTyrThrIleLeuLysAspSerArgPhe 552
DB 99 GATCTCTACCGATGATGTGGATACCCAGATGTGGACCATCTTAAAGGACAGCCGATTT 158
QY 553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572
DB 159 TTCCGTACTTGCACACAGCTGTGATGATGAGTGGACCATGCTGTGTTCAGGAAAC 218
QY 573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592
DB 219 ACACAAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTTTCAGATTTTCATGGCC 278
QY 593 TyrAspIleAlaCysAspArgTyrSerValLeuProArgProAspLeuHisHisaspVal 612
DB 279 TATGACATTCCTGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTC 338
QY 613 AsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPhe 632
DB 339 AACAGATTTGGCCATTCAGCAGTCTTACACACAGCACCATGATGTGTTCGGTGGTTTC 398
QY 633 AsnSerLeuLeuLeuSerAspIleLeuValPheThrSerSerGln--CysAspAlaHisA 652
DB 399 AATAGTCTCTCTCTCAGCCACATCTGTGATTCACCTCGGAACAGTGTGCAT--GGCAATC 457
QY 652 rgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlyS 672
DB 458 GGAGTGAACCCGCTTGTGTAGCAGCAGGACCTGGTATTTCGGTGTGTGTGGAACACAGGT 517
QY 672 erSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeu 689
DB 518 CGTCTCAGTGTATCTCTGTGGCGCTGGCAACTGATGAAACAGAGAAAGATTA 570

RESULT 25

AY418588
LOCUS
DEFINITION AY418588 3671 bp DNA linear GSS 17-DEC-2003
Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY418588

AY418588.1	GI:39774548
GSS.	
Pan troglodytes (chimpanzee)	
Pan troglodytes	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.	
1 (bases 1 to 3671)	
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,	
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,	
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,	
Adams,M.D. and Cargill,M.	
Inferring nonneutral evolution from human-chimp-mouse orthologous	
gene trios	
Science 302 (5652), 1960-1963 (2003)	
24671302	
2 (bases 1 to 3671)	
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,	
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,	
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,	
Adams,M.D. and Cargill,M.	
Direct Submission	
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,	
Rockville, MD 20850, USA	
These sequences were made by sequencing genomic exons and ordering	
them based on alignment.	
Location/Qualifiers	
source	
1..3671	
/organism="Pan troglodytes"	
/mol_type="genomic DNA"	
/db_xref="taxon:9598"	
<!.>.3671	
/genes="ATRN"	
/locus_tag="HCM6607"	
gene	
ORIGIN	
Alignment Scores:	
Pred. No.: 3,74e-98 Length: 3671	
Score: 135.00 Matches: 229	
Percent Similarity: 98.28% Conservative: 0	
Best Local Similarity: 98.28% Mismatches: 2	
Query Match: 9.45% Indels: 4	
DB: 29 Gaps: 0	
US-09-787-097-12 (1-1429) x AY418588 (1-3671)	
Qy 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyr 1177	
Db 2978 ACTCTTCATTGACTATCATAGTTACGTTTAGCTATCCAGGAAGATCGCTATTAC 3037	
Qy 1178 ThrAlaIleAsnPheValAlaThrProAspGlnAsnAraAspLeuAspMetPheIle 1197	
Db 3038 ACAGCTATTAAATTTTGCGGTACTCTCGACAACAAACAGGGATTGGACATGTTATC 3097	
Qy 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217	
Db 3098 AATGCCTCCAAGAATTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACC 3157	
Qy 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237	
Db 3158 CAGCTGGAGAGAGATGCGCTTGTTTTCGCAACCCAAACCAATTAAGAGAGTACAAAGATAGT 3217	
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257	
Db 3218 TTCCTAATGAGAAGTTGATTTCGCAACCCCAAATATCATCTTCTTGTTTATGTC 3277	
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPhe-SerGlnHisSerAsnPheM 1277	
Db 3278 AGTAATTTACCTGGCCCATCAAAATTCAGATGCGCTCTCTCNNN-AC-ACCAATTTTA 3335	
Qy 1277 etAspLeuValGlnPhePheValThrPhePheSerCysPheSerLeuLeuValAla 1297	
Db 3336 TGGACCTGGTACAGTTCTGCGACTTCTTCAGTGTCTTCCCTCTCTTCTGCTCGTGG 3395	

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-191200-001-b07&t3=2000-12-19&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 598.

FEATURES

Location/Qualifiers

1..598

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="EN0093"

/note="Organ: lung normal; Vector: puc18; Site_1: Smal;

Site_2: Smal; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.: 2,678-97 Length: 598

Score: 133.00 Matches: 167

Percent Similarity: 98.82% Conservative: 0

Best Local Similarity: 98.82% Mismatches: 1

Query Match: 9.31% Indels: 2

DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853006 (1-598)

QY 1095 GlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHis 1114

Db 6 GGCTTCTACGGTGTATCCCAATGAGGGAATGTGAGCAATGCAATGGGCAC 65

QY 1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134

Db 66 GGGTCTCTGTGCAACACCAACGCGCAAGTGTCTTGCCACCAACAGGGCGTCAAGGGG 125

QY 1135 AspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThr 1154

Db 126 GACGAGTGGCAGTATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACA 185

QY 1155 CysTyrTyrThrLeuLeuLeuLeuTyrGlnPheThrPheSerLeuSerGlnGluAsp 1174

Db 186 TGTATTATATCTCTCTTATTGACTATCATCTTACCTTTAGTCTATCCAGGAAGATGAT 245

QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAsp 1194

Db 246 CGGTATTACACGCTATCATTTTGTGGTACTCTCTGACGACAAACAGGCAATTGGAC 305

QY 1195 MetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIleAsnPheSer 1214

Db 306 ATGTTTCATCAATGCCTCCAGAGATTTCAACCTCAACATCACCTGGGCTGCCGATTTCTCA 365

QY 1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerLys-ThrAsnIleLysGlu 1234

Db 366 GTGGAACCCAGCTGGAGAGAGATGCTGTGTTTC-AAGAACCAACATTAAAGAGTA 424

QY 1234 rLysAspSerPheSerAsnGluLysPheAspPheArgHisProAsnIleThrPhePh 1254

Db 425 CAAAGATAGTTTCTCTAATAGAGAGTTTGATTTTCGCAACCAACCAATATCACTTCTT 484

QY 1254 eValTyrValSerAsnPheThrTip 1262

Db 485 TGTTCAGTCAGTAATTCACCTGG 509

RESULT 29

CA502816/c

LOCUS

DEFINITION

UI-CF-FNO-afg-d-10-0-UI-s1 UI-CF-FNO Homo sapiens cDNA clone

UI-CF-FNO-afg-d-10-0-UI 3', mRNA sequence.

CA502816

VERSION

CA502816.1 GI:24993770

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 739)

Authors

Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

889548

COMMENT

Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 16-237, >L1PA2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

Location/Qualifiers

1..739

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="UI-CF-FNO-afg-d-10-0-UI"

/tissue_type="Human Lung Epithelial cells"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-FNO"

/notes="Organ: Lung; Vector: p77T3-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

UI-CF-FNO is a subtracted cDNA library derived from two

normalized Human lung epithelial cell libraries (EN1 and

DUI) The library was subtracted according to according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:

bento-soares@uiowa.edu

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr to LPS 24h

TAG LIB=UI-CF-FNO

TAG_SEQ=CTGCTCAGGT

ORIGIN

Alignment Scores:

Pred. No.: 5,73e-94 Length: 739

Score: 129.00 Matches: 129
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.03% Indels: 0
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA502816 (1-739)

QY 1139 LeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThr 1158
 DB 622 CTATGTGAGTAGAATCGATACCAAGGAACCCCTCTCAGAGACATGTTATTATATCT 563
 QY 1159 LeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThr 1178
 DB 562 CTCTCTTATTGACATCATCAGTTTCACTTTAGTCTATCCAGGAAGATGCGCTATTACACA 503
 QY 1179 AlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsn 1198
 DB 502 GCTATCAATTTTGTGGCTACTCTCGACGAACAAACAGGGATTGGACATGTTCATCAAT 443
 QY 1199 AlaSerLysAsnPheAsnLeuAsnIleThrTTPAlaAlaSerPheSerAlaGlyThrGln 1218
 DB 442 GCCTCCAGAAATTTCAACCTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGACCCAG 383
 QY 1219 AlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPhe 1238
 DB 382 GCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTC 323
 QY 1239 SerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSer 1258
 DB 322 TCTAATGAGAGTTTGAATTTTGCACACACCCCAATATCACTTTCTTTTGTATGTCAGT 263
 QY 1259 AsnPheThrTrpProIleLysIleGln 1267
 DB 262 AATTTTCACTGGCCCATCAAAATTCAG 236

RESULT 30

CA895878 503 bp mRNA linear EST 20-DEC-2002
 B0195B01-5N NIA Mouse Neural Stem Cell (Differentiated) cDNA
 Library (Long) Mus musculus cDNA clone NIA:B0195B01 IMAGE:30102924
 5', mRNA sequence.

ACCESSION

CA895878

VERSION

CA895878.1 GI:27347427

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 503)

AUTHORS

Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.

TITLE

Systematic Analyses of NIA Mouse Neural Stem Cell (Differentiated)

JOURNAL

Unpublished (2002)

COMMENT

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 400, Baltimore, MD 21224-6820, USA

Email: cdna@gsun.grc.nia.nih.gov

Plate: B0195 row: B column: 01

Seq primer: -21M13 Reverse

High quality sequence stop: 503

POLYA=No.

Location/Qualifiers

1. 503

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CD1"

/lab host="DH10B"

/clone_lib="NIA Mouse Neural Stem Cell (Differentiated)
 cDNA Library (Long)"

/note="vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://gsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were
 obtained from Dr. Angelo L. Vescovi (Institute for Stem
 Cell Research, Italy). Double-stranded cDNAs were
 synthesized with an Oligo(dT) primer [Invitrogen:
 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from
 2.0 microgram of total RNA, treated with T4 DNA
 polymerase, and purified by ethanol-precipitation. The
 cDNAs were ligated to lone-linker L1-Sal4, purified by
 phenol/chloroform, and separated from free linkers by
 Centricon 100. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase
 (Takara) with a primer Sal4-S. The products were purified
 by phenol/chloroform and Centricon 100. The cDNAs were
 digested with SalI and NotI enzymes and cloned into
 SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
 coli host was transformed with the ligation mixture by the
 standard chemical method. The average insert size is about
 3.2 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
 Pred. No.: 1.86e-88 Length: 503
 Score: 122.00 Matches: 122
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.54% Indels: 0
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA895878 (1-503)

QY 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
 DB 138 CCGTGTGTTTCAAAACCAACATCAAGGAATACAAAGATAGCTTCTTAATCAGAAATTT 197
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrPro 1263
 DB 198 GATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTGAGTAATTCATTGGCCC 257
 QY 1264 IleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePhe 1283
 DB 258 ATCAAAATTCAGATTGCTTCTCCAGCAGACAGCACTTCATGAGCTGGTACAGTTCTTC 317
 QY 1284 ValThrPhePheSerCysPheLeuSerLeuLeuValAlaValValTrpLysIle 1303
 DB 318 GTGACTTTCTTCAGTTGTTTCTCTCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
 QY 1304 LysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMet 1323
 DB 378 AGCAGAGCTGTGGCATCCAGCGGAGAGCAACTTCTTCGGAGATGCAACAGATG 437
 QY 1324 AlaSerArgProPheAlaSerValAsnValAlaLeuLeuThrAspGluProProAsp 1343
 DB 438 GCCAGCGCCCTTTGCTTCTGTAACGTTGCTTTGGAACAGATGAAGAACTCTCTGAT 497
 QY 1344 LeuIle 1345
 DB 498 CTCATT 503

RESULT 31

CA320165

LOCUS

CA320165

DEFINITION

UI-M-FW0-cbx-g-11-0-UI-r1 NIH BMAP_FW0 Mus musculus cDNA clone

ACCESSION

CA320165

VERSION

CA320165.1 GI:24538289

686 bp mRNA linear EST 09-JUL-2003

IMAGE:6815748 5', mRNA sequence.

EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
1 (bases 1 to 686)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .686
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6815748"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP FWO"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is ACGGACACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

Program Coordinator:									
ORIGIN									
Alignment Scores:									
Pred. No.:	1.63e-87	Length:	686						
Score:	121.00	Matches:	134						
Percent Similarity:	99.26%	Conservative:	0						
Best Local Similarity:	99.26%	Mismatches:	0						
Query Match:	8.47%	Indels:	1						
DB:	14	Gaps:	0						
US-09-787-097-12 (1-1429) x CA320165 (1-686)									
QY	1296	ValAlaValValTtpLysIleLysGlnSerCysTrpAlaSerArgArgGluGln	1315						
Db	2	GTGCTCGAGTGGTCTCGAAGATCAACAGAGCTGTGGGCATCCAGCGGAGAGACAA	61						
QY	1316	LeuLeu-ArgGluMetGlnGlnMetalSerArgProPheAlaSerValAsnValAlaLe	1335						
Db	62	CTTCTTCGGGAGATGCAACAGATGGCCAGCGGCCCTTTGCTCTCTAAACGTTGCCCT	121						
QY	1335	uGluThrAspGluclupProAspLeuIleGlyGlySerIleLysThrValProLysPr	1355						
Db	122	GSAAACAGATGAAGAACTCTGATCTCATTTGGGGGAAGTATAAAGACCGTTCCTAAGCC	181						
QY	1355	cIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLe	1375						

Db 182 CATTGCGCTGGAGCCCTGTCTTGGTAACAAAGCCGAGTCCTCTCTGTATTCGTGAGGCT 241

Qy 1375 uProArgGlyLeuGlyClylleProProProdglyGlnSerGlyleuAlaValaSerAl 1395

Db 242 CCCTCGAGGACTGGGAGGAATCCCTCTCTGTCAGTCAGGTCCTGCTGTGGCCAGTGC 301

Qy 1395 aLeuValAspIleSerGlnGlnMetProIleValTyrlYsGluLysSerGlyAlaValAr 1415

Db 302 CCTGTGGCATTTCTCAGCAGATGCCAATAGTGACAAAGAGAAGTCAGGAGCTGTAAg 361

Qy 1415 gAsnArgLySgInGlnPrProProlaGlnProGlyThrCySille 1429

Db 362 AAACCGGAAGCAGCAGCGCGCTCACAGCCTGGAACCTGCATT 404

RESULT_32
BI089332 706 bp mRNA linear EST 20-JUN-2001
LOCUS 602853292F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4994735 5'
DEFINITION mRNA sequence.
ACCESSION BI089332
VERSION BI089332.1 GI:14507662
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NTH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 706)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11017 row: g column: 24
High quality sequence stop: 676.

FEATURES
Location/Qualifiers
source 1..706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4994735"
/cell_line="MGC36"
/lab_host="DH10B"
/clone_lib="NIH_MGC_10"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 7.02e-86 Length: 706
Score: 119.00 Matches: 180
Percent Similarity: 98.90% Conservative: 0
Best Local Similarity: 98.90% Mismatch: 1
Query Match: 8.33% Indels: 2
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BI089332 (1-706)

Qy 508 AspHisArgThrArgAlaLeuTyrlValHisGlyGlyrYsAlaPheSerAlaAsnLys 527

Db 32 GACCAATAGGACACAGCCCTATACGTTTCATGGTGGCTACAGGCTTCAGTGCCAATAAG 91

Qy 528 TyrArgLeuAlaAspAspLeuTyrlArgTyrAspValAspThrGlnMetTrpThrIleLeu 547

Db 92 TACC GGCTTG CAGATGATCTCTACCGATATGATGGGATACCGATGTGGACCAATCTT 151

Qy 548 LysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeu 567
 Db 152 AAGGACAGCCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGAACCATGCTG 211
 Qy 568 ValPhe-GlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSe 587
 Db 212 GTGTA-TGGAGGAACACACACACATGACACATCTATGAGCCATGGCCCAATGCTTCTC 270
 Qy 587 rSerAspPheMetAlaTyrAspIleAlaCysAspArgTyrSerValLeuProArgProAs 607
 Db 271 TTCAGATTTTCATGGCTATGACATTCCTGTGACCGCTGGTCACTGCTTCCAGACCTGA 330
 Qy 607 PleuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTy 627
 Db 331 TCTCCACCCATGATGTCACAGATTTGGCCATTCAGCAGCTTACACACAGACCATGTA 390
 Qy 627 rValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluG1 647
 Db 391 TGTGTTGCGTGGTTTCAATAGTCTCTCTCCTCAGCAGCATCTGGTATTACCTCGGAACA 450
 Qy 647 nCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysVa 667
 Db 451 GTGTGATGCGCATCGGAGTGAAGCGCTTGTCTAGCAGCAGGACCTGGGTATACGGTGTGT 510
 Qy 667 lTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluG1 687
 Db 511 GTGGACACAGGGTCTCTCAGTGTATCTCTGCGGCGCTGGCACTGATGACACAGAGA 570
 Qy 687 uIys 688
 Db 571 AAAG 574

RESULT 33
 LOCUS CD644584
 DEFINITION AGENCOURT_14539338 NTA Human H1 Embryonic Stem Cell cDNA Library
 (Long) Homo sapiens cDNA clone IMAGE:30421895 5', mRNA sequence.
 ACCESSION CD644584
 VERSION CD644584.1 GI:31816608
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 792)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
 cDNA Library Preparation: Yulan Piao and Minoru Ko
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC c lone distribution information
 can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDAM500 row: 9 column: 24
 High quality sequence stop: 681.
 location/Qualifiers
 1..792
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30421895"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NTA Human H1 Embryonic Stem Cell cDNA Library"

FEATURES

source

(Long)"
 /note=vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
 This is a long-transcript enriched cDNA library (Genome
 Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01
 cell line. Undifferentiated human ES cell line WA01/H1
 was obtained from WiCell Research Institute, Inc.,
 Madison, WI, cultured according to their instructions, on
 MEP feeders. They formed round colonies with defined edges
 and were positive for alkaline phosphatase, SSEA-4, OCT3,
 OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
 negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
 TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days
 after plating), the ES cells from 4 X 6cm dishes were
 treated with 1 mg/ml collagenase, type IV
 (Invitrogen/GIBCO) for 5-10 min and gently scraped off
 with 5 ml pipette. RNA was purified with TRIzol Reagent
 from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
 (2001). [PMID: 11544199] Double-stranded cDNAs were
 synthesized with an oligo(dT) primer [Invitrogen:
 5'-pGAGTAGTCTAGATCGGAGCGCGCCCTTTT-TTTT-3'] from
 3.4g of total RNA, precipitated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lone-linker LL-Sal4, purified by phenol/chloroform
 extraction, and separated from free linkers by
 Centricon-100 column. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase
 (Takara) with a primer Sal4-S for 25 cycles. The products
 were purified by phenol/chloroform extraction and
 Centricon-100 column. The cDNAs were digested with SalI
 and NotI enzymes and cloned into SalI/NotI site of
 pCMV-SPORT6 plasmid vector. The average insert size is
 about 3.6kb."

ORIGIN

Alignment Scores:
 Pred. No.: 7,85e-86 Length: 792
 Score: 119.00 Matches: 219
 Percent Similarity: 99.10% Conservative: 0
 Best Local Similarity: 99.10% Mismatches: 1
 Query Match: 8.33% Indels: 2
 DB: 14 Gaps: 0
 US-09-787-097-12 (1-1429) x CD644584 (1-792)
 Qy 161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
 Db 14 CTCATTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCATCATCTTTGCTACAGAG 73
 Qy 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
 Db 74 TGTAGTTGGACCAATTTATATGTTATGATGGGACTCAATTTATGACCGTAGTGTGT 133
 Qy 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
 Db 134 GCATTTAGTGGCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTGT 193
 Qy 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
 Db 194 GCCACATCAGGTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
 Qy 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
 Db 254 TTTAATATTACTTACAGTTTGTATGTTGTTCCAAATCTCTCAGCCGACGAGAGTGT 313
 Qy 261 LysIleSerAsnSerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyG1 280
 Db 314 AAGATCAGTAATAGACGCGA-TACTGTTCAATGTCAATGTTCTGAAACTGGAAGGTGA 372
 Qy 280 uIaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAs 300
 Db 373 AGCATGTGACATCTCTCACTGTACAGACAACTGTGTGTTTCTCTCATCGAGGATCTCGCA 432
 Qy 300 nSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVa 320

```

Db      433  TTCAAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGTCCTGGAATTCAGT 492
Qy      320  lProValProAlaSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPr 340
Db      493  TCCGTGACCACTAACCAAGTCATTTTGACCTCGAGAGGAATATTCCTAACTTAAGCTCCC 552
Qy      340  cArgAlaSerHisLysAlaValAlaValAsnGlyAsnIleMetTrpValValGlyGlyTyrMe 360
Db      553  CAGAGCATCTCATMAAGCTGTGGTCAATGGAACATTAATGTGGTGTGGAGGATATAT 612
Qy      360  tPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAsnLeuAlaSerArgGluTrpLe 380
Db      613  GTTCAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCT 672
Qy      380  u 380
Db      673  T 673

RESULT 34
BG673763 390 bp mRNA linear EST 04-FEB-2002
LOCUS 9010 ciliated epithelial cDNA cell library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.
ACCESSION BG673763
VERSION BG673763.2 GI:18490081
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 390)
AUTHORS Maiti,A.K., Jorissen,M. and Bouvagnet,P.
TITLE Isolation, in silico characterization and chromosomal localization
of a group of cDNAs from ciliated epithelial cells after in vitro
ciliogenesis
JOURNAL Genome Biol. 2 (7), RESEARCH0026 (2001)
MEDLINE 21407920
PubMed 11516339
COMMENT On May 1, 2001 this sequence version replaced gi:13919402.
Contact: Maiti AK
Laboratory de Genetic Molculaire Humaine, Faculty de Pharmacy
University Cl. Bernard,
8 Avenue Rockefeller, F69373 Lyon cedex, France
Tel: (33) 478 77 44 25
Fax: (33) 478 77 75 68
Email: amit.maiti@medecine.unige.ch
Seq primer: M13 Forward and reverse.
FEATURES
Location/Qualifiers
source
1..390
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Nasal biopsy"
/cell_type="cultured ciliated epithelial cells after in
vitro ciliogenesis"
/dev_stage="Adult"
/clone_lib="ciliated epithelial cDNA cell library"
/notes="Vector: PC82.1, cloned RT-PCR fragment from
isolated total RNA"

ORIGIN
Alignment Scores: 1.64e-84 Length: 390
Pred. No.: 117.00 Matches: 117
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 8.19% Indels: 0
Query Match: 12 Gaps: 0
DB:

US-09-787-097-12 (1-1429) x BG673763 (1-390)

Qy      285  ProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVal 304

```

```

Db      2  CTCACTGTACAGACAACCTGCGTTTCTCATCGAGGCATCTGCAATTCAGTGAATGC 61
Qy      305  ArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAla 324
Db      62  AGAGGATGCTCCTGCTTCTCAGACTGGCAGGTCCTGGATGTTCACTTCTGTACCAGCT 121
Qy      325  AsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHis 344
Db      122  AACCACTCATTTTGGACTCGAGAGGAATATCTTAACCTTAAGCTCCCCAGAGCATCTCAT 181
Qy      345  LysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSer 364
Db      182  AAAGCTGTGGTCAATGGAACATTAATGTGGTGTGGAGGATATATGTTCAACCACTCA 241
Qy      365  AspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArg 384
Db      242  GATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGT 301
Qy      385  SerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLys 401
Db      302  TCTGTGAACAAATGTTGTTGATATGATGTCATCTTGTGCATTATACAG 352

RESULT 35
BE259981 493 bp mRNA linear EST 26-OCT-2000
LOCUS 601148539F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163953 5',
DEFINITION mRNA sequence.
ACCESSION BE259981
VERSION BE259981.1 GI:9130910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 493)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM125 row: m column: 10
High quality sequence stop: 490.
FEATURES
Location/Qualifiers
source
1..493
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3163953"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 19"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores: 1.34e-83 Length: 493
Pred. No.: 116.00 Matches: 149
Score: 98.03% Conservative: 0
Percent Similarity: 98.03% Mismatches: 0
Best Local Similarity: 98.03%

```

Query Match: 8.12% Indels: 3
 DB: 10 Gaps: 0
 US-09-787-097-12 (1-1429) x BE259981 (1-493)

QY 414 ThrGlyAsnValThrLanGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
 DB 41 ACTGGGAATGTGCAATCAGTGTGAGAGTTTTCACATTCAATAGTCAATGAGTCAATGGTGTG 100
 QY 434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSer--AlaHisIle-Val 452
 DB 101.TTGACCCCTTAAGCAAGACGACGATGAGTGTGGCATCTCCGGCACACATGGTT 160
 QY 453 ThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGly 472
 DB 161 ACACCTGAAGAATGGCCGAGTGGTCACTGCTGGTCACTCTTTGGTCACTGGCCCTCTATGGA 220
 QY 473 TyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThr 492
 DB 221 TATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGATATATACACACC 280
 QY 493 GlnGlyAlaLeuValGlnGlyTyrGlyHisSerSerValTyrAspHisArgThrArg 512
 DB 281 CAGGGTGCCTTGTGCAAGGGGTTACGCCATAGCAGTGTTTACGACCATAGGACCGG 340
 QY 513 AlaLeuTyrValHisGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaSer 532
 DB 341 GGCCTATACGTTATGCTGGCTACAGGCTTTTCAGTGCCAAATAGTACCGGCTTGACAT 400
 QY 533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552
 DB 401 GATCTACCGATATGATGTGATACCCAGATGTGACCATCTCTTAGGACAGCCGATTT 460
 QY 553 PheArgTyrLeuHisThrAlaValLeuVal 562
 DB 461 TTCCCGTTACTTGACACAGCTGTGATAGT 490

RESULT 36
 BF880955 399 bp mRNA linear EST 17-JAN-2001
 LOCUS QV1-ET0181-021200-519-d10 ET0181 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION BF880955.1 GI:12271081
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 399)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,K.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 2020263
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV1&t2=QV1-ET0181-
 021200-519-d10&t3=2000-12-02&t4=1)
 Seq primer: puc 18 forward

High quality sequence start: 11
 High quality sequence stop: 399.
 Location/Qualifiers
 source 1..399
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev stage="Adult"
 /clone_lib="ET0181"
 /note="Organ: lung tumor; Vector: puc18; Site 1: SmaI;
 Site2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Alignment Scores: Length: 399
 Pred. NO.: 7.03e-83 Matches: 115
 Score: 115.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 8.05% Gaps: 0
 DB: 10

US-09-787-097-12 (1-1429) x BF880955 (1-399)

QY 1054 AsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsn 1073
 DB 27 AACTGGTCTTTCATTCTCTGCTCCAGTTCGCCATGCAACGGCCACAGTAATGATCAAT 86
 QY 1074 GlnSerIleCysGlnLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIle 1093
 DB 87 CAGAGCATCTGTGAGAAGTGTGAGAACTCTCACACAGGCAAGCACTGCGAGACCTGCATA 146
 QY 1094 SerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGly 1113
 DB 147 TCTGGCTTCTACCGTATCCCAATGAGGGAATGTGACCATGCAATGG 206
 QY 1114 HisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLys 1133
 DB 207 CACGGTCTCTGTGCAACCAACACACGGCAAGTCTTCTGCACACCAAGCGGTCAAG 266
 QY 1134 GlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGly 1153
 DB 267 GGGGACGAGTGGCAGTATGTGAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGA 326
 QY 1154 ThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSer 1168
 DB 327 ACATGTTACTATACTCTTCTTATTGACTATCATGTTCACTTTAGT 371

RESULT 37
 BF881492/c
 LOCUS UI-CF-EN1-adh-j-10-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
 DEFINITION
 ACCESSION BF881492
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 681)
 Ronald,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 McCray Lab

University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arranged by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 16-237, >LIPAZ*LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

```

1.681
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-adh-j-10-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pTV73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pTV73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)16 tail. The sequence tag for this library is CTGCTCAGGT. TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h TAG_LIB=UI-CF-EN1 TAG_SEQ=CTGCTCAGGT"

```

ORIGIN

Alignment Scores:	
Pred. No.:	1,199-82
Score:	115.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	8.0%
DR:	12
Length:	681
Matches:	115
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-787-097-12 (1-1429) x BM981492 (1-681)

Qy	1153	GlyThrCysTyrTyrThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGlu	1172
Db	580	GGAAATGTTATTATATCTCTCTTTATTGACTATCAGTTCACCTTTAGCTATCCCAAGAA	521
Qy	1173	AspAspArgTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAtcAsp	1192
Db	520	GATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAAACAGGGAT	461
Qy	1193	LeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSer	1212
Db	460	TTGACACATGTCATCAATGCCTCCAGAAATTTCAACCTCAACATCACCCTGGGCTGCCAGT	401
Qy	1213	PheSerAlaGlyThrGlnAlaGlyGluGlnMetProValValSerLysThrAsnIleLys	1232
Db	400	TTCTCAGTGTGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACAACATTAAAG	341

Qy	1233	GlutTyriyAspSerPheSerAsnGluuLyysPheaspPheargAsnHisProAsnIleThr	1235
Db	340	GAGTACAAAGATAGTTCTCTAATGAGAAGTTGATTTTCGCACCAACCAATATCACT	281
Qy	1253	PhePheValTyrValSerAsnPhethrTrrPproileLyysIleGln	1267
Db	280	TTCTTTGTTTATGTCAAGTAAATTTCACTGGCCCAACAAAATTCAG	236
RESULT 38			
EGS63729			
LOCUS	602584512F1 NIH MGC_76	674 bp	linear
DEFINITION	EST 10-APR-2001		
	602584512F1 NIH MGC_76 Homo sapiens cDNA clone IMAGE:4712114 5', mRNA sequence.		
ACCESSION	EGS63729		
VERSION	EGS63729		
KEYWORDS	EGS63729.1	GI:13571381	
SOURCE	EST.		
	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1. (bases 1 to 674)		
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .		
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)		
TITLE	Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT			

Email: cgapbs-@email.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L12CM1553 row: h column: 03
High quality sequence stop: 674.

FEATURES

```

1. .574
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/cJones="IMAGE:4712114"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 76"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site:1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTTCAGGCCGAGCGCGACATG-dt (30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.88
.kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
cal). Note: this is a NIH MGC library."

```

ORIGIN:

Alignment Scores:		
Pred. No.:	4, 91a-81	574
Score:	113.00	208
Percent Similarity:	97.20%	0
Best Local Similarity:	97.20%	3
Query Match:	7.91%	6
DB:	12	0
TS-09-787-097-12	(1-1429) x	RG563729 (1-674)

US-09-787-097-12 (1-1429) x BG563729 (1-674)

QY 857 ValGlyIeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876

DB 2 GTCGGCCTTCGGAAGATCAATGTGCTCCPACTGGTGGTGGGAAGATATGTCCCCATTACA 61

QY 877 AsnSerIeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896

DB 62 AATAGTTTACTACAGTGGATCCCGTCTGAGCCCACTGATCTGGATTCTGTGGAATTATA 121

```

QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db 122 TCAGAACCCAGGACTCGGGAGCTGAAGGTGCAACCTGCATCAACCCATCAATGGTAGT 181
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db 182 GTCTGTGAAGGCTCTCAACACACAGTCTAAGCAGTCCGACACCATGTGCTTGAGG 241
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db 242 ACAGCATGTGGAGATTGCACCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAG 301
QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPhe--GlyGlnCysMetGluT 976
Db 302 CAGTGTGTGAGCTCCATGCTATGTGGCTCTT-CCCTTTGTGGCCAGTGTATGGAA 360
QY 976 TrpTyrThrMetSerThrCysProGluAsnCys-SerGlyTyrCysThrCysSerHis 995
Db 361 GGTATACGATGAGCACCTGCCCTCTGAAATGG-TTCAGGCTACTGTACCTGTAGTCA 419
QY 996 CysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLys 1015
Db 420 TCGTTGGACCAACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 479
QY 1016 CysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsn 1035
Db 480 TGCATAGAGGGTTCCTATTAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGGAAT 539
QY 1036 PheTyrProGlnProLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrp 1055
Db 540 TTCTATCCACAGCC-CTGCTCAATTCAGCATGTGTCTAGAGGACAGACATACAACTGG 598
QY 1056 SerPheIleHisCysProAlaCysGlnCysAsnGly 1067
Db 599 TCTTTCATTCAGTCCAGCTTGCCTATGCAACGGC 634

RESULT 39
AA350293/c
LOCUS EST57512 Infant brain Homo sapiens mRNA linear EST 21-APR-1997
DEFINITION AA350293
ACCESSION AA350293
VERSION AA350293.1 GI:2002620
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
AUTHORS 3,400 expressed sequence tags identify diversity of transcripts
TITLE from human brain
JOURNAL Nat. Genet. 4, 256-267 (1993)
MEDLINE 93364420
PUBMED 8358434
COMMENT Other ESTs: EST57511 THCL168226
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 400
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):151062"
/db_xref="taxon:9606"
/sex="female"
FEATURES
source
1. .671

```

```

/dev_stage="infant"
/Clone_lib="Infant brain"
/notes="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI"

ORIGIN
Alignment Scores: 1.23e-79 Length: 400
Pred. NO.: 111.00 Matches: 111
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 7.77% Gaps: 0
DB: 9

US-09-787-097-12 (1-1429) x AA350293 (1-400)

QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 334 CCTGCAACACACAGTGTGAAGCAGTCCGACACCATGTGCTTGGAGCAGCATGTGGA 275
QY 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 274 GATTGCACCAACGCGNAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGG 215
QY 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 214 TCCAATGCCTATGTGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 155
QY 981 ThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnPro 1000
Db 154 ACCTGCCCCCTGAAATTTGTCAGCTACTGTGTACCTGTAGTCATTGCTTGGAGCAACA 95
QY 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGlyLys 1020
Db 94 GGTGTGGCTGTGTACTGATCCAGCAATATCGGCAAGGGAATGATAGAGGGTTCC 35
QY 1021 TyrLysGlyProValLysMetProSerGlnAla 1031
Db 34 TATAAGGACCAAGTGAAGTGCCTTCGCAAGCC 2

RESULT 40
AA1818764/c
LOCUS AA1818764
DEFINITION AA1818764
ACCESSION AA1818764
VERSION AA1818764.1 GI:5437843
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 671)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2162 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 411.
Location/Qualifiers
1. .671

```


/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:244606"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP Uti"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

ORIGIN

Alignment Scores:
Pred. No.: 1,32e-78 Length: 671
Score: 110.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.70% Indels: 0
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AI818764 (1-671)

QY 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177
DB 547 ACTCTTCTTATTGACTATCAGTTCACTTTCAGTCTATCCCGAAGATGATCGCTATTAC 488
QY 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
DB 487 ACAGTATCAATTTTGTGGTACTCTCGACGACAAACACAGGATTGGACATGTTTCATC 428
QY 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
DB 427 AATGCTCCACGAATTTCAACTCAACATCACTGGCTGGCAGTTTCTCAGTCGAACC 368
QY 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
DB 367 CAGGCTGGAGAGAGATGCTGTTTTCATAAAACCAACATTAGGAGTACAAAGATAGT 308
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
DB 307 TTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCAATATACATTTCTTTGTTTATGTC 248
QY 1258 SerAsnPheThrTrpProIleLysIleGln 1267
DB 247 AGTAATTTCACTGGCCCATCAAAATTCAG 218

RESULT 41

AL048842
LOCUS
DEFINITION
DKFZp434G2118_r1 434 (synonym: htes3) Homo sapiens cDNA clone

AL048842

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 502)

Ottawaelder, B., Oberwarter, B., Mewes, H.W., Gassenhuber, J. and

Wiemann, S.

EST (Ottawaelder, et al.)

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by MedGenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project. No sl sequence

available.

This clone (DKFZp434G2118) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..502
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434G2118"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI"

ORIGIN

Alignment Scores:
Pred. No.: 4.16e-77 Length: 502
Score: 108.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.56% Indels: 0
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AL048842 (1-502)

QY 159 ThrTrpLeuIleGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAla 178
DB 2 ACGTGGCTCATTCAGACGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGCT 61
QY 179 ThrGluCysSerTrpAspHisIleuTyrValTyrAspGlyAsnSerIleTyrAlaProLeu 198
DB 62 ACAGAGTGTAGTTGGGACCAATTATATGTTATGATGGGACTCAATTTATGACCGCTA 121
QY 199 ValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGlu 218
DB 122 GTTGTGCTATTAGTGGCTCATTTGCTCAGAGAGATGGCAATCAGACTGTCCCTGAG 181
QY 219 ValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaIleTyrAsnLeu 238
DB 182 GTTGTGGCCACATCAGGTTATGCTTGTGTCATTTTTTTAGTCATGCTGCTTAAATTTG 241
QY 239 ThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGly 258
DB 242 ACTGGATTTAATATTACTTACAGTTTGTATGATGTGTCCTCAATTAATCTGCTCAGCCGAGGA 301
QY 259 GluCysLysIleSerAsnSerSer 266
DB 302 GAGTGTAAAGATCAGTAATAGCAGC 325

RESULT 42

BF963698/c

LOCUS

DEFINITION

PM4-NN1204-221200-001-g03 NN1204 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF963698

VERSION

BF963698.1

KEYWORDS

EST.

1 GI:12380973

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 484)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, F.S., Sucher, P., Jongeneel, C.V.,

O'Hare, M.J.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4tt2-PM4-NN1204-221200-001-903st3-2000-12-22&f4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 350.
FEATURES Location/Qualifiers
 1..484
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="NN1204"
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
ORIGIN
 Alignment Scores:
 Pred. No.: 7,038-74 Length: 484
 Score: 104.00 Matches: 135
 Percent Similarity: 98.54% Conservative: 0
 Best Local Similarity: 98.54% Mismatches: 1
 Query Match: 7.28% Indels: 2
 DB: 10 Gaps: 0
 US-09-787-097-12 (1-1429) x BF963698 (1-484)
 QY 1025 VallysMetProSer:GlnAlaProThrGlyAsnPhetYrProGlnProLeuLeuAsnSer 1044
 Db 482 GTGAAGATGCCCTCGCAAGCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCC 423
 QY 1045 SerMetCysLeuGluAspSerArgYrAsnTrpSerPheIleHisCysProAlaCysGln 1064
 Db 422 AGCATGTGTCAGAGGACAGACATACCACTGGTCTTTTCATTCCTACTGTCCAGCTTGCCAA 363
 QY 1065 CysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084
 Db 362 TCGAAGCGCCACAGTAATGATCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 303
 QY 1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheYrGlyAspProThrAsnGlyGly 1104
 Db 302 ACAGGCAAGCACTGGAGACCTGCAATCTGGCTCTTACGGTGATCCCAATGGAGGG 243
 QY 1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuLeuCysAsnThrGlyLys 1124
 Db 242 AAATGTGAGCATGCAAGTGCATGAGGACGGCTCTGTGCAACCAACACGGGCAAG 183
 QY 1125 CysPheCysThrThr-LysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAs 1144
 Db 182 TGCCTTCGCACNC-CAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAA 124
 QY 1144 nArgYrGlnGlyAsnProLeuAspGlyThrCysYrYrThrLeuLeu 1160
 Db 123 TCGATCCAGGAACCTCTTCAGAGGAACATGTTATTACTCTTCT 75
RESULT 43
LOCUS AI372809/c 383 bp mRNA linear EST 12-JAN-1999

DEFINITION EST175353 Infant brain, Bento Soares Homo sapiens cDNA clone
FEATURES FUHIBU1, mRNA sequence.
ACCESSION AI372809
VERSION AI372809.1 GI:4152675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (Bases 1 to 383)
AUTHORS Adams,M.D., Kexlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,D.M., Fitzhugh,W.M., Fritchman,J.B., Geoghagen,N.S.,
 Glodok,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.P., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Pannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
 EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
 EST175339 EST17534
 Contact: HGI (Human Gene Index)
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-0200
 Fax: (301)-838-0208
 Email: hgi@icr.org.
FEATURES Location/Qualifiers
 1..383
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="FUHIBU1"
 /lab_host="E. coli DH5-alpha"
 /clone_lib="Infant brain, Bento Soares"
 /note="Vector: BA, M13-derived; Site 1: HindIII; Site 2:
 NotI; The infant brain library, constructed by Bento
 Soares, Columbia University, was oligo-(dT) primed and
 directionally cloned into an M13-derived plasmid using
 total brain mRNA from a 72-day old human female afflicted
 with spinal muscular atrophy."
ORIGIN
 Alignment Scores:
 Pred. No.: 6,338-70 Length: 383
 Score: 99.00 Matches: 99
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.93% Indels: 0
 DB: 9 Gaps: 0
 US-09-787-097-12 (1-1429) x AI372809 (1-383)
 QY 822 LeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIle 841
 Db 297 CTTTGGCTCTCTTACAAACCAGAGAGGAGGATTTGCTTACGACGTCGGAATA 238
 QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLys 861

Db 237 ATGAGTCATCTCAGAGCATGTCAGAGCTACCTTAACCCCATGGTGGCTTCGGAAG 178

Qy 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881

Db 177 ATCAATGTGCTCTACTGTGTGGAGATATGCCCCATTTACAAATAGTTTACTACAG 118

Qy 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901

Db 117 TGGATGCCGTCTGAGCCAGTGATGCTGGATCTCTGGAATTTTATCAGAACCCAGTACT 58

Qy 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920

Db 57 CGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCACTCAATGTAGTGTCTGTGAAGG 1

RESULT 44

CA874419

LOCUS

DEFINITION K0932C03-SN NIA Mouse Neural Stem Cell (Undifferentiated) cDNA

Library (long) Mus musculus cDNA clone NIA:K0932C03 IMAGE:30087290

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.

1 (bases 1 to 567)

Systematic Analyses of NIA Mouse Neural Stem Cell

(Undifferentiated) cDNA Library (long)

Unpublished (2002)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@ngs.grc.nia.nih.gov

Plate: K0932 row: C column: 03

Seq primer: -21M13 Reverse

High quality sequence stop: 567

POLYA=No.

FEATURES

source

1..567

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CD1"

/db_xref="taxon:10090"

/db_xref="taxon:10090"

/clone="NIA:K0932C03 IMAGE:30087290"

/dev_stage="Adult"

/lab_host="DH10B"

/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)

cDNA Library (long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://ngs.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Angelo L. Vescovi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGAGGCGGCCCTTTTCTTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loxe-linker IL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into

ORIGIN

Alignment Scores:

Pred. No.: 9,28e-70 Length: 567

Score: 99.00 Matches: 99

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 6.93% Indels: 0

DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA874419 (1-567)

Qy 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861

Db 240 ATGCAATCATCTCAAGATATGTCAGAGTCATCTGACTCCATGGTTCGTTTCGGAAG 299

Qy 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881

Db 300 ATCAATGTGCTCTACTGTGTGGAGGATATGCTCCATTCAAAATAGTTTGTCTGAG 359

Qy 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901

Db 360 TGGATGCCATCTGAGCCAGTGATGCTGGCTTCTGTGGATCTTGTGAGAGCCTAGTACT 419

Qy 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921

Db 420 CGGGATTAAGGCTGCACCTGCATCAACCTCTCAATGGCAGCGTCTGTGAAGGCCT 479

Qy 922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940

Db 480 GCAAAACACAGTGCAGCAGTCCGACACCATGTGCCCTGCGGACAGCGTGTGGC 536

RESULT 45

BF313142

LOCUS

DEFINITION BF313142 NIH MGC_19 Homo sapiens cDNA clone IMAGE:4126959 5',

mRNA sequence.

ACCESSION

VERSION

BF313142.1 GI:11261090

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 953)

NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L10M1017 row: j column: 16

High quality sequence stop: 612.

Location/Qualifiers

1..953

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4126959"

/tissue types="neuroblastoma"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 19"

/note="Organ: Brain; Vector: pOTB7; Site 1: XhoI; Site 2:

ECCKI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1,54e-69 Length: 953
Score: 99.00 Matches: 154
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 1
Query Match: 6.93% Indels: 1
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF313142 (1-953)

QY 414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
Db 61 ACTGGAAATGTGACCAATGATGATGAGATTTTCATTCATATGATGATGATGATG 120
QY 434 LeuThrProLysAlaLysGluGlnTrpAlaValValGlyHisSerAlaHisIleValThr 453
Db 121 TTGACCCCTAAGCAAGGAGCAGTATGAGTGGTGGCACTCTGCACACATTTTACA 180
QY 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTrpGlyTyr 473
Db 181 CTGAAGAATGGCCGAGTGTGATGCTGTCATCTTTGGTCACTGCTCTCTATGATAT 240
QY 474 IleSerAsnValGlnGluTrpAspLeuLysAsnThrTrpSerIleLeuHisThrGln 493
Db 241 ATAAGCAATGTGACCAATATGATTTGGATAGACACATGAGATATATACACCCAG 300
QY 494 GlyAlaLeuValGlnGlyGlyTyrGlyHisSerValTyrAspHisArgThrArgAla 513
Db 301 GGTGCCCTTGTGCAAGGGGGTTCACGCCATAGCAGTGTGTTACGACCATAGGACCCAG-GCC 359
QY 514 LeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAsp 533
Db 360 CTATAGCTTCATGTTGGTTCAGAGCTTTCAGTGCATATAGTACCGCTTGCAGATGAT 419
QY 534 LeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePhe 553
Db 420 CTCCTACCGATATGATGATGATACCCAGATGTGACCAATCTTAAAGACACGCCGATTTTC 479
QY 554 ArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuVal 568
Db 480 CGTTACTTGCACACAGCTGTGATGTGATGAGTGAACCATGCTGGTG 524

RESULT 46
LOCUS BF853681
DEFINITION MR2-EN0093-211200-003-a05 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853681
VERSION BF853681.1 GI:12241425
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 595)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR2&t2=MR2-EN0093-211200-003-a05&t3=2000-12-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.

FEATURES
source

1..595
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/notes="Organ: lung_normal; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 6,29e-69 Length: 595
Score: 98.00 Matches: 157
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 1
Query Match: 6.86% Indels: 2
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853681 (1-595)

QY 1095 GlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHis 1114
Db 7 GGCTTCTACGGTGATCCCAATGAGGGAATGTCACCATGCAATGGCAATGGGCAC 66
QY 1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134
Db 67 GCCTCTCTGTGCAACACACACACGCGCAAGTCTCTGCACCAAGGCGCTCAAGGG 126
QY 1135 AspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThr 1154
Db 127 GACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGAAC 186
QY 1155 CysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAsp 1174
Db 187 TGTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCACGAGATGAT 246
QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu-As 1194
Db 247 CGCTATTACACGCTATCAATTTTGTGGTACTCTCTGACGAACAAACAGGGA-TTGGGA 305
QY 1194 pMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSe 1214
Db 306 CATGTTCAATGCTCTCAAGATTTTCACTCAACATCAGCTGGGCTGCCAGTTTCTC 365
QY 1214 rAlaGlyThrGlnAlaGlyGluMetProValValSerLysThrAsnIleLysGluTyr 1234
Db 366 AGCTGAACCCAGCTGGAGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGAGTA 425
QY 1234 rLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThr 1252
Db 426 CAAAGATAGTTTCTCTAATGAGAGATTGATTTTCGAACACCAACCAATATCACT 480

RESULT 47

CB525640
LOCUS
DEFINITION UI-M-FY0-cfd-j-16-0-UI_r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6848081 5', mRNA sequence.
ACCESSION
VERSION CB525640
KEYWORDS
SOURCE CB525640.1 GI:29359097
ORGANISM Mus musculus (house mouse)
MUS MUSCULUS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 565)
NIH-MGC <http://mgc.nci.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uicwa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
Location/Qualifiers
1..565
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6848081"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is ACCGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

Alignment Scores:
Pred. No.: 1..62e-66 Length: 565
Score: 95.00 Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.65% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB525640 (1-565)

QY 1335 LeuGluThrAspGluGluProProAspLeuIleGlySerIleValProlys 1354
Db 2 TTGGAACAGATGAAGACCTCTGATCTCAATGGGGGAAGTATAAGACCGCTTCTTAAG 61
QY 1355 ProfileAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValarg 1374
Db 62 CCCATGCCCTGGAGCCCTGCTTTGGTAAACAAAGCGCGAGTCTCTGTTATCTGTGAGG 121
QY 1375 LeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSer 1394

Db 122 ATCCCTCGAGGACTGGAGGAATCCCTCTCTCTGCTCAGGTCAGTCTGCTGTCAGCT 181
QY 1395 AlaLeuValAspIleSerGlnGlnMetProIleValTylySerGlyAlaVal 1414
Db 182 GCCTCTGCTGACATTTCTCAGCAGATGCCAATAGTGTACAAAGAGAGTCAAGAGCTGTA 241
QY 1415 AtgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db 242 AGAAACCGAGAGCAGCGCGCTGCAGAGCTGGAACTGCATT 286
RESULT 48
AI372810/c
LOCUS
DEFINITION EST175334 Infant brain, Bento Soares Homo sapiens cDNA clone
FUIBUI, mRNA sequence.
ACCESSION
VERSION AI372810.1 GI:4152676
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 320)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Falanges, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.D., Kunsch, C., HungJun, J., Li, H., Meissner, P.S., Olsen, H.,
Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
Fraser, C.M. and Venter, J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT
Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
EST175339 EST17534
Contact: HGI (Human Gene Index)
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-0200
Fax: (301)-838-0208
Email: hgi@tigr.org
FEATURES
Location/Qualifiers
1..320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FUIBUI"
/lab_host="E. coli DH5-alpha"
/clone_lib="Infant brain, Bento Soares"
/note="Vector: BA, M13-derived; Site 1: HindIII; Site 2:
NotI; The infant brain library, constructed by Bento
Soares, Columbia University, was oligo-(dT) primed and
directionally cloned into an M13-derived plasmid using
total brain mRNA from a 72-day old human female afflicted
with spinal muscular atrophy."

ORIGIN
Alignment Scores:
Pred. No.: 3.88e-65 Length: 320

Pred. No.: 5.98e-62 Length: 281
 Score: 89.00 Matches: 89
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.23% Indels: 0
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BG015062 (1-281)

QY 588 SerAspPheMetAlaTyAspIleAlaCysAspArgTrpSerValLeuProArgProAsp 607
 Db 279 TCAGATTTCATGGCTGATGACATTGCTGACCGCTGGTCAGTCTCCAGACCTGAT 220
 QY 608 LeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyr 627
 Db 219 CTCACCAATGATGACACAGATTTGGCCATTACGAGCTTTACACACAGACCATGTAT 160
 QY 628 ValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerGluGln 647
 Db 159 GTGTTCCGTGTTTCAATAGTCTCTCTCCTCAGCGACATCTCTGATTACCTCGGAACAG 100
 QY 648 CysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysVal 667
 Db 99 TGTGATGGCATCGAGTGAAGCCGCTGTATTAGCAGCAGGACCTGGTATTTCGGTGTGTG 40
 QY 668 TrpAenThrGlySerSerGlnCysIle 676
 Db 39 TGGACACAGGGTGTCTCAGTGATC 13

RESULT 51

BF853021/c 483 bp mRNA linear EST 16-JAN-2001
 LOCUS MR2-EN0093-191200-001-905 EN0093 Homo sapiens cDNA, mRNA sequence.

DEFINITION BF853021

ACCESSION BF853021.1 GI:12240869

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 483)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,

Goldman,G.H., Carvalho,A.F., Matekuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR2&t2=MR2-EN0093-

191200-001-905&t3=2000-12-19&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 483.

Location/Qualifiers

1..483

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="EN0093"

FEATURES

source

1..483

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="EN0093"

ORIGIN

Alignment Scores:
 Pred. No.: 1.02e-61 Length: 483
 Score: 89.00 Matches: 89
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.23% Indels: 0
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853021 (1-483)

QY 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
 Db 446 GAGGTAGAAAATCATACCAAGAAACCTCTCAGAGGAACATGTTATTATCTCTCTT 387
 QY 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
 Db 386 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATTCGCTATTACACAGCTATC 327
 QY 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
 Db 326 AATTTTGTGGCTACTCTCTGACGAACAAACAGGGATTTGGACATGTTCAATGCTCTC 267
 QY 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
 Db 266 AAGATTTCACCTCAACATCACCTCGGCTGCAGTTTCTCAGCTGGAACCCAGGCTGGA 207
 QY 1221 GluGluMetProValValSerLysThr 1229
 Db 206 GAAGAGATGCTGCTGTTGTTTCAAAACCC 180

RESULT 52

AA350292

LOCUS

DEFINITION

AA350292.1

ACCESSION

AA350292.1

VERSION

AA350292.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 396)

Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.

3,400 expressed sequence tags identify diversity of transcripts

from human brain

Nat. Genet. 4, 256-267 (1993)

93364420

MEDLINE

PUBMED

8358434

COMMENT

Other_ESTs: EST57512.JHC168885

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavage@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (http://www.tigr.org/tdb/hgi/hgi.html)

Seq primer: M13-21.

Location/Qualifiers

1..396

/organism="Homo sapiens"

/mol_type="mRNA"


```

/db_xref="ATCC (inhost):151062"
/db_xref="taxon:9606"
/dev="female"
/dev stage="infant"
/clone_lib="infant brain"
/clone="Organ: brain; Vector: latmid BA; Site_1: HindIII; Site_2: NotI"

ORIGIN
Alignment Scores:
Pred. No.: 2,266-59 Length: 396
Score: 86.00 Matches: 123
Percent Similarity: 98.40% Conservative: 0
Best Local Similarity: 98.40% Mismatches: 1
Query Match: 6.02% Indels: 2
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AA350292 (1-396)
QY 480 TTTATGTTTCGATGAGTATATATACACCCAGGCGCCCTTGCGACGG 60
DB 1 TATGATTTCGATGAGTATATATACACCCAGGCGCCCTTGCGACGG 60
QY 500 GYTYTGlyHisSerValTyrAspHisArgThrArgAlaLeuTyrValHis-GlyGI 519
DB 61 GGTACGGCCATAGCAGTGTTCAGCACCATAGCAGCAGGCGCTATACGTTNA-TGGTGG 119
QY 519 YTYrlyAlaPheSerAlaAsnlyTyArgLeuAlaAspPheLeuTyrArgTyrAspVa 539
DB 120 CTACAGGCTTTTCAGTGCCAAATAGTACCGCTTCAGATGATCTCTACCGATATGATGT 179
QY 539 lAspThrGlnMetTrrpThrIleLeuLyAspSerArgPheArgTyrLeuHisThrAl 559
DB 180 GGATACCCAGATGTGGACCATCTTAAAGACAGCGCATTTTCCTTACTTGCACACAGC 239
QY 559 avallleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMe 579
DB 240 TGTGATGATGAGTGAACCATCTCGTGTGGAGGAACACACACATATGACATCTAT 299
QY 579 tSerHisGlyAlaLyCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspAr 599
DB 300 GAGCCATGCGGCCAAATGCTTCTTCAGATTTCATGCGCTATGACATGTGCTGTGACCG 359
QY 599 gTrpSerValLeu 603
DB 360 CTGGTCAGTGCCT 372

RESULT 53
CB522182
LOCUS
DEFINITION
UI-N-GH0-cep-c-09-0-UI-r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE:6842530 5', mRNA sequence.
CB522182
VERSION
CB522182.1 GI:29355537
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

```

Seq primer: pyX-5.
Location/Qualifiers
1. 758
/organism="Mus musculus"
/mol type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6842530"
/tissue_type="Whole brain"
/dev stage="1, 5, and 15 days newborn"
/lab host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GH0"
/clone="Organ: Brain; Vector: pyX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAATCGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 1,786-57 Length: 758
Score: 84.00 Matches: 84
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.88% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB522182 (1-758)
QY 1346 GlyGlySerIleLeuValProLyProIleAlaLeuGluProCysPheGlyAsnLys 1365
DB 15 GGGGGAGTATATAGACCCCTTCTTANGCCCATTCCTCGAGCCCTGCTTTGGTAACAA 74
QY 1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyIleProPro 1385
DB 75 GCGCAGTCTCTCTGTTATTCGAGGCTCCCTCGAGGACTGGAGGAATCCCTCTCTCT 134
QY 1386 GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPro 1405
DB 135 GGTCACTCAGGCTCTGCTGTGCCAGTGCCTGGTGACATTTCTCAGCAGATGCCATA 194
QY 1406 ValTyrLysGlyLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro 1425
DB 195 GTGTACAGGAGAGTCAAGAGTGTATAGAAACCGAAGCAGCAGCGGCTGCACAGCT 254
QY 1426 GlyThrCysIle 1429
DB 255 GGAACCTGCATT 266

RESULT 54
BU531905
LOCUS
DEFINITION
AGENCOURT 10205588 NIH_MGC_126 Homo sapiens cDNA clone
IMAGE:6556819 5', mRNA sequence.
BU531905
VERSION
BU531905.1 GI:22842346
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 778)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
 Email: c9pbbs-remail.nih.gov
 Tissue Procurement: NCI
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LICM2723 row: n column: 19
 High quality sequence stop: 480.
FEATURES Location/Qualifiers
 source
 1..778
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6556819"
 /tissue_type="mixed (pool of 40 RNAs)"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_lib="NIH MGC 126"
 /note="Vector: pDNR-LIB; Site 1: sfii (ggccattatggcc); Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGTATACGACGAGTGGCCATTCAGCGCGG-3' and 5'-ATTCAGTGGCGGCGGCGGCGCATG-dt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-1 kb size fraction (other fractions present in NIH MGC 127 and NIH MGC 128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,838-57 Length: 778
 Score: 84.00 Matches: 84
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.88% Indels: 0
 DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BUS31905 (1-778)

QY 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
 DB 4 GCTACTCTCGAGCAACAAACAGGATTTGACATGTTTCATCAATGCTCCCAAGATTC 63

QY 1204 AsnLeuAsnIleThrTrrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMet 1223
 DB 64 AACCTCAACATCACTGGGCTGCCAGTTCTCAGCTGGAAACCAGCTGGAGAGAGATG 123

QY 1224 ProValValSerLysThrAsnIleLysGluTyrlLysAspSerPheSerAsnGluLysPhe 1243
 DB 124 CCTGTTGTTTCAAAACCAACATTAAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 183

QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrlValSerAsnPheThrTrpPro 1263
 DB 184 GATTTCGCAACACCAACCAATATCACTTCTTTGTTTATGTCAGTAATTTACCTGGCCCC 243

1264 IleLysIleGln 1267
 DB 244 ATCAAAATTCAG 255

RESULT 55
 BI004303

LOCUS BI004303 484 bp mRNA linear EST 13-JUN-2001
DEFINITION PMO-HN0078-280201-010-h11 HN0078 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI004303
VERSION BI004303.1 GI:14408377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 484)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Congenel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&t2=PMO-HN0078-280201-010-h11&t3=2001-02-28&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 18
 High quality sequence stop: 483.
FEATURES Location/Qualifiers
 source
 1..484
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HN0078"
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from CRESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Alignment Scores:
 Pred. No.: 4,816-56 Length: 484
 Score: 82.00 Matches: 124
 Percent Similarity: 98.41% Conservative: 0
 Best Local Similarity: 98.41% Mismatches: 1
 Query Match: 5.74% Indels: 2
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BI004303 (1-484)

QY 1183 ValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsn 1202
 DB 75 GTGGCTACTCTCGAGCAACAAACAGGATTTGACATGTTTCATCAATGCTCCCAAGAT 134

QY 1203 PheAsnLeuAsnIleThrTrrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGlu 1222
 DB 135 TTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCAGGCTGGAGAGAG 194

QY 1223 MetProVal-ValSerLysThrAsnIleLysGluTyrlLysAspSerPheSerAsnGluTy 1242
 DB 195 ATGCCTGC-TGTTTCAAAACCAACCAATTAAGGAGTACAAAGATAGTTTCTTCTAATGAGAA 253

```

QY 1242 sPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTr 1262
Db 254 GTTGGATTTCGCACACCCACCAATATCACTTCTTTGTTATGTCAGTAATTCACCTG 313
QY 1262 pProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPh 1282
Db 314 GOCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTATGCGACCTGGTACACTT 373
QY 1282 ePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLy 1302
Db 374 CTTCGGACTTCTTCAGTGTGTTCTCTCTTCTGCTCTGGTGGCTGCTGTGGTTGGAA 433
QY 1302 sIleLysGlnSerCys 1307
Db 434 GATCAACAAGTGTGT 449

RESULT 56
CB520374
LOCUS CB520374 840 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GIO-csi-1-18-0-UI.r1 NIH_BMAP_GIO Mus musculus cDNA clone
IMAGE:6840067 5', mRNA sequence.
ACCESSION CB520374
VERSION CB520374.1 GI:29353729
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1. (bases 1 to 840)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
COMMENT NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This genome was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..840
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840067"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GIO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACCGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES
source
1..840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3763690"
/db_xref="taxon:9606"
/clone="IMAGE:194640"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 5,33e-55 Length: 840
Score: 81.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.67% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB520374 (1-840)

QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
Db 374 ATCAATCATCTCAAGATATGTCACAGCTCAGCTCAGCTCCATGGGTCTGTTCTGGAAG 433
QY 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
Db 434 ATCAATGTGTCTTACTGGTCTGGGAGGATATGCTCCATTCACAAATAGTTTGTGCAG 493
QY 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
Db 494 TGGATGCCATCTCAGCCAGTGAATGCTGCTTCTGTGGGATCTTGTGCAGAGCTAGTACT 553
QY 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
Db 554 CGGGGATTAAGGCTGCAACCTGTCATCAACCTCTCAATGGCAGCGCTCTGTGAAGGCT 613
QY 922 Ala 922
Db 614 GCA 616

RESULT 57
R87650
LOCUS R87650 416 bp mRNA linear EST 16-AUG-1995
DEFINITION YP89G01.s1 Soares fetal liver spleen INFUS Homo sapiens cDNA clone
IMAGE:194640 3', mRNA sequence.
ACCESSION R87650
VERSION R87650.1 GI:946473
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 416)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1984
High quality sequence stops: 273
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1984 Std Error: 0.00
Seq primer: Promega -2lml3
High quality sequence stop: 273.
Location/Qualifiers
1..416
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3763690"
/db_xref="taxon:9606"
/clone="IMAGE:194640"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"

```

/clone lib="Soares fetal liver spleen INFLS"
 /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
 1st strand cDNA was primed with a Pac I - Oligo(dT) primer
 15' AACGGAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3',
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Patima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 1.74e-54 Length: 416
 Score: 80.00 Matches: 80
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.60% Indels: 0
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x R87660 (1-416)

QY 753 MetTyrCysAsnLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGln 772
 Db 1 ATGTACTACTGTAAACAGAACACAGCTGCAGAGCTGCGCCCTGGACAGAACTGCCAG 60
 QY 773 TtpGluProArgAsnGlnCysileAlaLeuProGluAsnileCysGlyleGlyTyr 792
 Db 61 TGGGAGCCCGGAATCAGAGTGCTATGCCCTGCCGGAATAATCTGGGCATTCGCTGG 120
 QY 793 HisLeuValGlyAsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAla 812
 Db 121 CATTTGGTTGGAAATCATGTTTGAATTAATCTACTGCCAAGAGAGATTAATGACAATGCT 180
 QY 813 LysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysLysVal 832
 Db 181 AAATTTGTTCTGTAGAACCAACATGCCCTTTTGGCTTCTCTACCAACCCAGAGAGTA 240

RESULT 58

BF854487/c
 LOCUS BF854487 567 bp mRNA linear EST 16-JAN-2001
 DEFINITION MR2-EN0093-271200-005-a04 EN0093 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF854487

VERSION BF854487.1 GI:124242231

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 567)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR2&t2=MR2-EN0093-

271200-005-a04&t3=2000-12-27&t4=1)

Seq primer: puc 18 forward
 High quality sequence stop: 566.

FEATURES

source

1. 567
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="EN0093"
 /note="Organ: lung normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 2.35e-54 Length: 567
 Score: 80.00 Matches: 166
 Percent Similarity: 97.65% Conservative: 0
 Best Local Similarity: 97.65% Mismatches: 2
 Query Match: 5.60% Indels: 4
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF854487 (1-567)

QY 1096 PheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAla 1115
 Db 567 TTCTACGGTGATCCACCAATGGAGGAATGTGAGCCATGCAAGTGCATGGCAGCG 508
 QY 1116 SerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAsp 1135
 Db 507 TCTCTGTGCAACCAACACGGGCAAGTGCTTCTGCACCACCAAGGCGCTCAAGGGGAC 448
 QY 1136 GlnCysGlnLeuCysGluValGluAsnArgTyrGlnGly-AsnProLeuArgGlyThrCy 1155
 Db 447 GAGTGCCAGCTATGTGAGGTAGAAATCGATACCAAGA-AAACCTCTCAGAGGAACATG 389
 QY 1155 sTyrTyrThrLeuLeuileAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspAr 1175
 Db 388 TTATTACTCTCTTATTGACTATCATGTTACCTTTAGTCTATCCCGAGGAGATGATCG 329
 QY 1175 gTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMe 1195
 Db 328 CTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAACAGGAGATTGGACAT 269
 QY 1195 tPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTyrAlaAlaSerPheSerAl 1215
 Db 268 GTTCATCATGCTCCCAAGAAATTCACCTCAACATCCTGGCTGCCAGTTCTCAGC 209
 QY 1215 aGlyThrGlnAlaGlyGluMetProValValSer-LysThrAsnIleLysGluTyrL 1235
 Db 208 TGGAAACCAGGCTGGAGAGAGATCCCTGTGTTCAAAAGCC-AACATTAAAGAGATACA 150
 QY 1235 ysAspSerPheSerAsnGluLysPheAspPheArgenHisProAsnIleThrPhePheV 1255
 Db 149 AAGATAGTTTCTCTAAATGAGAAGTTGATTTCGCACCAACCAACCAATATCACTTCTTG 90
 QY 1255 alTyrValSerAsnPheThrTyrPro 1263
 Db 89 TTTATGTCAGTAATTTTCCCTGGCCC 64

RESULT 59

AM579586

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AM579586 427 bp mRNA linear EST 16-MAR-2000
 DEFINITION MRO-HT0209-100100-103-e09 HT0209 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM579586
 VERSION AM579586.1 GI:7254635
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 427)
 HCGP <http://www.ludwig.org.br/ORESTES>.
 The FAPESP/LICR Human Cancer Genome Project
 Unpublished (1999)
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-HT0209-100100-103-e09&t3=2000-01-10&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 427.
 Location/Qualifiers
 1..427
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0209"
 /note="Organ: head neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 FEATURES
 source
 1..427
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0209"
 /note="Organ: head neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,15E-53 Length: 427
 Score: 79.00 Matches: 115
 Percent Similarity: 98.29% Conservative: 0
 Best Local Similarity: 98.29% Mismatches: 1
 Query Match: 5.53% Indels: 2
 DB: 10 Gaps: 0
 US-09-787-097-12 (1-1429) x AW579586 (1-427)
 QY 434 LeuThrProValAlaValGlnTyrAlaValGlyHisSerAlaHisLeuValThr 453
 DB 64 TTGACCCCTAAGCAAGGAGGAGTATGCGTGGTGGGACCTCTGCACATTTGTACA 123
 QY 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473
 DB 124 CTGAGAATGGCGAGTGTGTCATCTGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 183
 QY 474 IleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGln 493
 DB 184 ATAAGCAATGTGCAGAAATATCATTTGGATAAGAACACATGGAGTATTATACACACCCAG 243
 QY 494 GlyAlaLeuValGlnGlyGlyTyrGlyHisSerValTyrAspHisArgThrArgala 513
 DB 244 GGTGCCCTTGTGCAGAGGGGTTACGGCCATAGCATAGTGTTCAGCACATAGGACGAGGG- 302
 QY 514 -LeuTyrValHisGlyGlyTyrIleAlaPheSerAlaAsnLysTyrArgLeuAlaAspAs 533
 DB 303 CCTATACGTTTCATGTGGTGTACAAAGCTTTTCAGTCCCAATAGTACCGGCTTGCAGATGA 362
 QY 533 pLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAsp 549
 DB 363 TCTTACCATATGATGTGGATACCCAGATGATGGACCATTTCTTAAGGAC 411
 RESULT 60
 BI202540
 LOCUS BI202540 434 bp mRNA linear EST 29-APR-2002

DEFINITION UI-M-CGOp-brh-h-01-0-UI-s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
 UI-M-CGOp-brh-h-01-0-UI 3', mRNA sequence.
 ACCESSION BI202540
 VERSION 1
 KEYWORDS B1202540.1 GI:14668512
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 434)
 Bonaudo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 COMMENT
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mest@mail.nih.gov
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 The tissue for this library was contributed by Dr. Xin-Yuan Fu,
 Yale University School of Medicine
 Seq primer: M13 Forward
 POLYA=No.
 Location/Qualifiers
 1..434
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CGOp-brh-h-01-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_Ret4_S2"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP Ret4 S2 library is a subtracted library,
 ultimately derived from mouse retina tissue libraries at
 various stages of development. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.uiowa.edu. The tissue
 for this library was contributed by Dr. Xin-Yuan Fu, Yale
 University School of Medicine
 TAG_SEQ=None found"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,16E-51 Length: 434
 Score: 76.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.32% Indels: 0
 DB: 12 Gaps: 0
 US-09-787-097-12 (1-1429) x BI202540 (1-434)
 QY 123 GlyTrpValGlyGluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSer 142
 DB 45 GGCTGGGTGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAACTGGCTCTCT 104
 QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIle 162
 DB 105 GGATTTGTAAACAGATGACCTGGGAATTATAAATATAGCAAGATGCACATGGCTCAT 164

Qy 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
 Db 165 GAAGACAGCCAAATAGATAATGAGACTTCCTCCCTTCAACCAATTTGCTACAGATGAGC 224
 Qy 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
 Db 225 TGGGACCATTATATGTTATGATGGGACTCAATCTACGCCCTCTG 272

RESULT 61
 AW120845
 LOCUS
 DEFINITION UI-M-BH2.3-aca-e-06-0-UI.s1 NIH BMAP M.S3.3 Mus musculus cDNA clone
 UI-M-BH2.3-aca-e-06-0-UI.3', mRNA sequence.
 ACCESSION AW120845
 VERSION AW120845.1 GI:6096178
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)

889548
 97044477
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
 available by the means that is soon to be determined. When NIH
 determines the means for distribution of the BMAP cDNA clones, this
 record will be updated accordingly when that means is determined.
 Seq primer: M13 Forward
 POLYA=No.

FEATURES

Location/Qualifiers
 1..435
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH2.3-aca-e-06-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M.S3.3"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site: Not I; Site 2: Eco RI; The
 NIH BMAP M.S3.3 Library is a subtracted library of a
 series, ultimately derived from a mixture of individually
 tagged, normalized libraries from ten regions of the mouse
 brain (cerebellum, brain stems, olfactory bulbs,
 hypothalamus, cortex, amygdala, basal ganglia, pineal
 gland, striatum, hippocampus) after a series of
 subtractions to reduce the representation of cDNAs from
 which ESTs had already been generated. The following
 serially subtracted libraries were generated in this
 process: NIH BMAP M.S3.3, NIH BMAP M.S2, NIH BMAP M.S1.
 The subtracted library (NIH BMAP M.S3.3) was constructed
 as follows: PCR amplified cDNA inserts from NIH BMAP M.S2
 clones from which 3' ESTs had been derived was used as a
 driver in a hybridization with the NIH BMAP M.S2 library
 in the form of single-stranded circles. The remaining
 single-stranded circles (subtracted library) was purified
 by hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into DH10B
 bacteria (Life technologies) to generate the

NIH BMAP_M.S3.3 library. This procedure has been
 previously described (Bonaldo, Lennon and Soares, Genome
 Research 6:791-806, 1996)
 TAG TISSUE=hypothalamus
 TAG LIB=NIH BMAP_M.S3.3
 TAG_SEQ=CGGTA"

ORIGIN

Alignment Scores: 3.17e-51 Length: 435
 Pred. No.: 76.00 Matches: 76
 Score: 76.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.32% Indels: 0
 DB: 10 Gaps: 0
 US-09-787-097-12 (1-1429) x AW120845 (1-435)
 Qy 123 GlyTrrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSer 142
 Db 44 GCGTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAACTGCTCTTCT 103
 Qy 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeu 162
 Db 104 GGATTTCTAACAGATGGACCTCGGAATTATAATATAAGACGAAGTGCACATGGCTCA 163
 Qy 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
 Db 164 GAAGACAGCCAAATAGATAATGAGACTTCCTCCCTTCAACCAATTTGCTACAGATGAGC 223
 Qy 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
 Db 224 TGGGACCATTATATGTTATGATGGGACTCAATCTACGCCCTCTG 271

RESULT 62

BI202684
 LOCUS
 DEFINITION UI-M-CG0p-brk-h-05-0-UI.s1 NIH BMAP Rat4.S2 Mus musculus cDNA clone
 UI-M-CG0p-brk-h-05-0-UI.3', mRNA sequence.

ACCSSION BI202684
 VERSION BI202684.1 GI:14668656
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery

Genome Res. 6 (9), 791-806 (1996)
 97044477

PUBMED

COMMENT

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov

Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Benco Scores
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 The tissue for this library was contributed by Dr. Xin-Yuan Fu,
 Yale University School of Medicine
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 Location/Qualifiers
 1..435

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0p-brk-h-05-0-UI"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP Ret4 S2"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP Ret4 S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu. The tissue
for this library was contributed by Dr. Xin-Yuan Fu, Yale
University School of Medicine
TAG_SEQ=None found"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      3,17e-51      Length:      435
Score:          76.00      Matches:      76
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      5.32%      Indels:      0
DB:              12      Gaps:      0

```

US-09-787-097-12 (1-1429) x B1202684 (1-435)

```

QY 123 GlyTnpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSer 142
DB 45 GGTGGTGGGGGAGAGAAAGCCAGACACTCGGGGGCGGCTTCAGACTAACGGCTCTCT 104
QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTnpLeulle 162
DB 105 GGTATTGTACAGATGAGACTCGGGAATTATAATATAGACGAGTGCACATGGCTCATT 164
QY 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
DB 165 GAAGGACACCCCAATAGATAATAGACTTCGCTTCAACCACTTTTGTACAGAAATGTAGC 224
QY 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
DB 225 TGGGACCAATTATATGTTTATGATGGGACTCAATCTACGACCTCTG 272

```

RESULT 63

```

AW124828
LOCUS
DEFINITION
UI-M-BH2.1-apk-d-10-0-UI.s1 NIH BMAP M.S3.1 Mus musculus cDNA clone
UI-M-BH2.1-apk-d-10-0-UI 3', mRNA sequence.
ACCESSION
AW124828
VERSION
AW124828.1 GI:6100358
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 488)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mES@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

```

Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
POLYA-No.

FEATURES

Location/Qualifiers

```

1..488
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH2.1-apk-d-10-0-UI"
/dev stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP M.S3.1"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP M.S3.1 library is a subtracted library of a
series, ultimately derived from a mixture of individually
tagged, normalized libraries from ten regions of the mouse
brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus) after a series of
subtractions to reduce the representation of cDNAs from
which ESTs had already been generated. The following
serially subtracted libraries were generated in this
process: NIH BMAP M.S3.1, NIH BMAP M.S2, NIH BMAP M.S1.
The subtracted library (NIH BMAP M.S3.1) was constructed
as follows: PCR amplified cDNA inserts from NIH BMAP M.S2
clones from which 3' ESTs had been derived was used as a
driver in a hybridization with the NIH BMAP M.S2 library
in the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the
NIH BMAP M.S3.1 library. This procedure has been
previously described (Bonaldo, Lennon and Soares, Genome
Research 6:791-806, 1996)
TAG TISSUE=amygdala
TAG LIB=NIH_BMAP_M.S3.1
TAG_SEQ=GTGAG"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      3,55e-51      Length:      488
Score:          76.00      Matches:      76
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      5.32%      Indels:      0
DB:              10      Gaps:      0
US-09-787-097-12 (1-1429) x AW124828 (1-488)
QY 123 GlyTnpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSer 142
DB 44 GGTGGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAACTGGCTTCT 103
QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTnpLeulle 162
DB 104 GGATTTGTAAACAGATGGACCTGGGAATTATAATATAAGCAAGTGCACATGGCTCATT 163
QY 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
DB 164 GAAGGACACCCCAATAGATAATAGACTTCGCTTCAACCACTTTTGTACAGAAATGTAGC 223
QY 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
DB 224 TGGGACCAATTATATGTTTATGATGGGACTCAATCTACGACCTCTG 271
RESULT 64
BE952846

```


LOCUS BE952846 566 bp mRNA linear EST 04-OCT-2000
 DEFINITION UI-M-BH4-azk-g-05-0-UI.s1 NIH BMAP_M S5 Mus musculus cDNA clone
 UI-M-BH4-azk-g-05-0-UI 3', mRNA sequence.
 ACCESSION BE952846
 VERSION BE952846.1 GI:10593734
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 566)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 Seq primer: M13 Forward
 POLYA=No.

FEATURES

source

Location/Qualifiers

1..566
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH4-azk-g-05-0-UI"
 /dev_stages="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M S5"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M_S5 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged,
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. For a detailed description of the library from
 which this clone was derived, please visit our web site
 at brainest.eng.uiowa.edu.
 TAG_SEQ=None found"

ORIGIN

Alignment Scores:
 Pred. No.: 4,1e-51 Length: 566
 Score: 76.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5,32% Indels: 0
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE952846 (1-566)

QY 123 GlyTpvAlGhGlnCysGlnHisCysGlyArgPheArgLeuThrGlySer 142
 Db 44 GCCTGGTGGGAGCAATGCCAGCACTGCGGGGCGCTTCAGACTAACTGGCTCTCT 103
 QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysThrLysCysThrTrpLeu 162

Db 104 GGATTTGTACAGATGACCTGGGAATTATAAATAGACGAAGTGCACATGGCTCAT 163
 QY 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
 Db 164 GAAGGACAGCCAAATAGAAATATGACATTCGCTTCACCACTTTGCTACAGATGAC 223
 QY 183 TtpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
 Db 224 TGGGACCAATTATATGTTTATGATGGGACTCAATCTAGCCACTCTG 271

RESULT 65

BE980188

LOCUS

DEFINITION

UI-M-BG2-bcq-f-12-0-UI.s1 NIH BMAP MSC S1 Mus musculus cDNA clone

UI-M-BG2-bcq-f-12-0-UI 3', mRNA sequence.

ACCESSION

BE980188

VERSION

BE980188.1 GI:10647894

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 626)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 Seq primer: M13 Forward
 POLYA=No.

FEATURES

source

Location/Qualifiers

1..626
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BG2-bcq-f-12-0-UI"
 /dev_stages="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP MSC S1"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_MSC_S1 library is a subtracted library derived
 from mouse spinal
 cord tissue. For a detailed description of the library
 from which this clone was derived, please visit our web
 site at brainest.eng.uiowa.edu.
 TAG_SEQ=None found"

ORIGIN

Alignment Scores:
 Pred. No.: 4,52e-51 Length: 626
 Score: 76.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5,32% Indels: 0
 DB: 10 Gaps: 0

```

US-09-787-097-12 (1-1429) x BE980188 (1-626)

QY 123 GlyTTPValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSer 142
Db 44 GGCTGGGTGGGAGCAATGCCAGCACTCGCGGGGCCCTTCAGACTAACTGGCTCTCT 103
QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuLe 162
Db 104 GGATTGTGAACAGATGGACCTGGCAATTATTAATAAGCAAGTGCACATGGCTCAAT 163
QY 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
Db 164 GAAGACACGCCAAATAGATAATAGACTTCGCTTCAACCACTTTTGCTACAGATGTAGC 223
QY 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
Db 224 TGGGACCACTTATATGTTTATGATGGGACTCACTACGACCTCTG 271

RESULT 66
BE671400/c
LOCUS
DEFINITION
7e50d11.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3285909 3'
similar to TR_060295 O60295 KIAA0548 PROTEIN ;contains L1.b1 L1
repetitive element ;, mRNA sequence.
ACCESSION
BE671400
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 445)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 397.
FEATURES
Location/Qualifiers
1..445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3285909"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pTVT3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI-CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.: 2,1e-50 Length: 445
Score: 75.00 Matches: 75
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.25% Indels: 0
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE671400 (1-445)

QY 1193 LeuAspMetPheIleAsnAlaSerIysAsnPheAsnLeuAsnIleThrTrpAlaLaseR 1212
Db 445 TTGGCATGTTTCATCAATGCCTCCAGAAATTTCAACCTCAACATCACCTGGCTGCCAGT 386
QY 1213 PheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerIysThrAsnIleLys 1232
Db 385 TTCTCAGCTGGACCCAGCGCTGGAGAGAGATGCCGTGTTGTTTCAAAAACCAACATTAAG 326
QY 1233 GluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThr 1252
Db 325 GAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATTTTTCGCAACCCACCAATATCACT 266
QY 1253 PhePheValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
Db 265 TTCTTTGTTTATGTCAGTAATTTACCTGCGCCATCAAAATTCAG 221

RESULT 67
AI076479/c
LOCUS
DEFINITION
OZ28B04.x1 Soares total fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:1676623 3' similar to contains L1.b1 L1 repetitive element ;,
mRNA sequence.
ACCESSION
AI076479
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 464)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1450 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 444.
FEATURES
Location/Qualifiers
1..464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1676623"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus_Nb2HF8_9w"
/note="Vector: pTVT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5',
TGTTACCAATCGAGTGGAGCGGCGCTTAATTTTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTVT3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.: 1,41e-49 Length: 464
Score: 74.00 Matches: 74
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 5.18% Indels: 0
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AI076479 (1-464)

QY 1194 AspMetPheIleAsnAlaSerLysAsnPheAsnLeuAenlierTTPAlaAlaSerPhe 1213
DB 463 GACATGTTCAATCAATGCTCCAGAAATTTCACTCACATCAGTGGCTGCCAGTTTC 404

QY 1214 SerAlaGlyThrGlnAlaGlyGluMetProValSerLysThrAsnIleLysGlu 1233
DB 403 TCAGTGAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGAG 344

QY 1234 TyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhe 1253
DB 343 TACAAGATAGTTTCTTAATGAGAAGTTTGATTTTGCACACCCCAAAATATCACTTTC 284

QY 1254 PheValTyrValSerAsnPheThrTTPProIleLysIleGln 1267
DB 283 TTGTTTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAG 242

RESULT 68
BF549645
LOCUS
DEFINITION UI-R-C2-nf-e-05-0-UI.r1 UI-R-C2 Rattus norvegicus cDNA clone
ACCESSION BF549645
VERSION BF549645.1 GI:11659375
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 474)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1776329 The following
repetitive elements were found in this cDNA sequence: 439-465,
>POLY A#Simple repeat
Seq primer: M13 Forward.
Location/Qualifiers
1. .474
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2-nf-e-05-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C2
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day

embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

ORIGIN

Alignment Scores: 9.33e-49 Length: 474
Pred. No.: 73.00 Matches: 73
Score: 73.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 5.11% Gaps: 0
DB: 10

US-09-787-097-12 (1-1429) x BF549645 (1-474)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuATGLysIleAsnValSerTTPCys 868
DB 20 TCCAGCTCACTGACTCCATGGTGGTTCGGAAGATCAATGATCTTACTCTGTGC 79

QY 869 TTPGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
DB 80 TGGAGGATATGTCCTCCATTCACAAATAGTTTGTCTGAGTGGATGCACTGAGCCAGT 139

QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
DB 140 GATGCCGGCTCTCTGGGATCTTGTGAGAGCTTAGTACTCGGGGAGTGAAGGCTGCAACC 199

QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
DB 200 TGCATCACCCACTCAATGGAAGTGTGTGAAAGGCGCT 238

RESULT 69
BE865031
LOCUS
DEFINITION UI-M-BH1-amt-f-11-0-UI.r1 NIH BMAP M.S2 Mus musculus cDNA clone
ACCESSION BE865031
VERSION BE865031.1 GI:10386668
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 535)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd, Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose

availability will be considered under appropriate and limited

collaborative arrangements

Seq primer: M3 Reverse:

Location/Qualifiers

FEATURES

source

1. .535

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH1-anc-f-11-0-UI"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH BMAP M S2"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not 1; Site 2: Eco RI; The

NIH BMAP M S2 library is a subtracted library derived from

derived from a mixture of normalized libraries from ten

regions of the mouse brain (cerebellum, brain stems,

olfactory bulbs, hypothalamus, cortex, amygdala, basal

ganglia, pineal gland, striatum, hippocampus). The driver

used for subtraction consisted of a pool of 5,000 clones

from the NIH BMAP M S1 library and a pool of 2,000 clones

obtained from non-normalized and normalized mouse brain

spinal cord libraries."

ORIGIN

Alignment Scores:

Pred. No.: 1 84e-45 Length: 535

Score: 69.00 Matches: 125

Percent Similarity: 98.43% Conservatives: 0

Best Local Similarity: 98.43% Mismatches: 1

Query Match: 4.83% Indels: 2

DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE865031 (1-535)

QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102

DB 1 CTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGCACCGCACTAAT 60

QY 1103 GlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThr 1122

DB 61 GGAGGCAATGTCAGCCATCGAAGTCAGATGGGACCATCTACTGTGCACACCAACACC 120

QY 1123 GlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluVal 1142

DB 121 GCGAAGTGTCTCTGATACCAACCAAGGTGTCAAGGGGACGAGTGCAGCTATGTGAGGTA 180

QY 1143 GluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuLeuAs 1162

DB 181 GAAATCGATACACGGGAACCTCTCAA-AGGAACATGCTACTATACCTCTTCATTGA 239

QY 1162 pTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPh 1182

DB 240 CTATCAGTTCACTTTAGCTGTCTCCAGGAGACGACCGCTACTACACAGCCATCACTT 299

QY 1182 eValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAs 1202

DB 300 TGTGCTACTCTGTGATGACAAACAGGAGTTTGACATGTTCAATGCTCCCAAAA 359

QY 1202 nPheAsnLeuAsnIleThr 1208

DB 360 CTTCAACCTCAACATCACC 378

RESULT 70

AI818750/c

LOCUS

DEFINITION

AI818750 589 bp mRNA linear EST 07-MAR-2000

similar to TR:060295 O60295 KIAA0548 PROTEIN ;contains L1.t3 L1

repetitive element ;, mRNA sequence.

ACCESSION AI818750

VERSION AI818750.1 GI:5437829

KEYWORDS

EST.

Homo sapiens (human)

Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapsb@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2056 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 399.

Location/Qualifiers

1. .589

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2424580"

/tissue_type="well-differentiated endometrial

adenocarcinoma, 7 pooled tumors"

/lab_host="DH10B"

/clone_lib="NCI CGAP Ut1"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 1.75 kb. Life Technologies catalog #:

11538-014"

ORIGIN

Alignment Scores:

Pred. No.: 2 02e-45 Length: 589

Score: 69.00 Matches: 112

Percent Similarity: 98.25% Conservatives: 0

Best Local Similarity: 98.25% Mismatches: 1

Query Match: 4.83% Indels: 2

DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AI818750 (1-589)

QY 1155 CysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAsp 1174

DB 556 TGTATTATATCTCTTTTATTGACTATATGATTCAGTTCACCTTCAGTCTATCCACGAGATGAT 497

QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgPheLeuAsp 1194

DB 496 CGCTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAAACAGGAGTTGGAC 437

QY 1195 MetPheIleAsn-AlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSe 1214

DB 436 ATGTTTCATCA-TGCTCTCAAGAAATTTTCACTTCAACATCAACATCAGCTGGCGTCCGAGTTTC 378

QY 1214 rAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyr 1234

DB 377 AGCTGGAACCCAGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 318

QY 1234 rLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePh 1254

DB 317 CAAAGATAGTTTCTCTTAATGAGAGAGTTGATTTTCGCAACACCCACCAATATCATTCTT 258

QY 1254 eValTyrValSerAsnPheThrTrpProIleLysIleGln 1267

DB 257 TGTTTATGTCAGTAATTTTCACTGGCCCAATCAAAATTCAG 218


```

/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30545215"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GV0"
/notes="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator."

```

ORIGIN

Alignment Scores:
 Pred. No.: 6,586-45 Length: 529
 Score: 68.00 Matches: 68
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.76% Indels: 0
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CD804170 (1-292)

```

QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTropValGlyLeuArgLys 861
Db 87 ATGCAATCATCTCAAGTAATGTCAGCTCATCTGACTCCAGTGGTGGTTCGGAAG 146
QY 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
Db 147 ATCAATGTCTTACTGTTGCTGGGAGGATATGCTCCATTACAAATAGTTTGTGCAG 206
QY 882 TrpMetProSerGluProSerAspLaglyPheCysGlyIleLeuSerGluProSerThr 901
Db 207 TGGATGCCATCTGAGCCCATGATGCTGCTTCTGTGGGATCTTGTGAGCCCTAGTACT 266
QY 902 AAGGlyLeuLysAlaAlaThrCys 909
Db 267 CGGGGATTAAGGCTGCAACCTGC 290

```

```

RESULT 73
BM537406
LOCUS
DEFINITION
  BM537406 529 bp mRNA linear EST 20-FEB-2002
  clone ha82a01 5', mRNA sequence.
ACCESSION
  BM537406
VERSION
  BM537406.1 GI:18817590
KEYWORDS
  EST.
SOURCE
  Canis familiaris (dog)
  Canis familiaris
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
  1 (bases 1 to 529)
  O'Shaughnessy,A.L., McCombie,W.R., Baker,J.P., Balija,V.,
  Cunniss,D., Dedhia,N.N., de la Bastide,W., Katzenberger,F.,
  King,L., Kirchoff,K.A., Miller,B., Muller,S., Nascimento,L.U.,
  Palmer,L., Santos,L., Shah,R.S., Spiegel,L.A., Zutavern,T.,
  Preston,R. and Hannon,G.J.
  Expressed sequence tags from Canis familiaris (dog) (2002)
  Unpublished (2002)
AUTHORS
  Contact: W. Richard McCombie
  Lita Aurenberg Hazen Genome Sequencing Center
  Cold Spring Harbor Laboratory

```

```

PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ha82 row: a column: 01
Seq primer: -21M13UnivRev
High quality sequence stop: 529.
Location/Qualifiers
  1..529

```

FEATURES

```

Source
  /organism="Canis familiaris"
  /mol_type="mRNA"
  /db_xref="taxon:9615"
  /clone="ha82a01"
  /tissue_type="testes"
  /clone_lib="Canis cDNAs from testes cells"
  /notes="Vector: lambda Zap II; The library was produced by
  Greg Hannon and Raymond Preston (Cold Spring Harbor
  Laboratory). This library is oligo(dT) primed using
  stragene zap cDNA synthesis kit. It was made from dog
  testes. Please contact Greg Hannon (hannon@cshl.org) with
  any library related inquiries."

```

ORIGIN

Alignment Scores:
 Pred. No.: 1,176-44 Length: 529
 Score: 68.00 Matches: 114
 Percent Similarity: 98.28% Conservative: 0
 Best Local Similarity: 98.28% Mismatches: 1
 Query Match: 4.76% Indels: 2
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM537406 (1-529)

```

QY 1037 TyrProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db 128 TACCACAGCCCTTTCATTCACATGTGCTGGAGGACAGACATACACTGCTCT 187
QY 1057 PheIleHisCysProIaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db 188 TTCATTCAATGTCCAGTTCAGTCCCAATGGCCACAGCAAAATGCAATTAATCAGATATC 247
QY 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db 248 TGTGAGAGTGTGAGAACCTGACCACAGGACAGACTGCGACCTGCATATCTGGCTTC 307
QY 1097 TyrGlyAspProThrAsnGlyGlyLys-CysGlnProCysLysCysAsnGlyHisAlaSe 1116
Db 308 TATGTTGATCCCAACCAATGG-GGGAAAGATGTGAGCCATGCAATGGCATGCCTC 366
QY 1116 rLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspG 1136
Db 367 TCTCTGTAAACAAACACGGCAAGTCTTGTGCACCACCAAGGGCGTCAAGGGGACCA 426
QY 1136 uCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeu 1151
Db 427 GTCCAGCTGTGTGAGTGGAAAATCGATACCAGGGAATCTCTC 472

```

RESULT 74

```

BM569805/c
LOCUS
DEFINITION
  BM569805 Sus Scrofa library (scac) Sus scrofa cDNA clone
  scac00291.b.13 3prim, mRNA sequence.
ACCESSION
  BM569805
VERSION
  BM569805.1 GI:37979704
KEYWORDS
  EST.
SOURCE
  Sus scrofa
  Sus scrofa (pig)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
  1 (bases 1 to 530)
  Bonnet,A., Tosser-Klopp,G., Benne,P., Cabau,C., Villegier,S.,
  Soares,M., Bonaldo,F. and Hately,F.
  A Pig Normalised Multi-Tissue cDNA Library

```